In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

3/17/65
b

# COUNTABLE MARKOV CHAINS
# WITH AN APPLICATION TO QUEUEING THEORY

A THESIS

Presented to

The Faculty of the Graduate Division

by

Ray Collins Owens, Jr.

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Mathematics

Georgia Institute of Technology

March, 1966

COUNTABLE MARKOV CHAINS

WITH AN APPLICATION TO QUEUEING THEORY

Approved:

Chairman

Date approved by Chairman: April 27, 1966

## ACKNOWLEDGMENTS

I am deeply indebted to my thesis advisor, Dr. Eric R. Immel, for his suggestions, guidance, and helpfulness, both in the preparation of this thesis and in my graduate studies. I also wish to thank Dr. James W. Walker and Dr. Harold A. Gersch for their help in preparing this thesis.

For the privilege of attending the Georgia Institute of Technology I owe a special thanks to my parents and to the University System of Georgia.

Finally, thanks are due Mrs. Betty Sims for her excellent typing of the thesis.

# TABLE OF CONTENTS


Page

ACKNOWLEDGMENTS. . . . . . . . . . . . . . . . . . . . . . . . . . ii

Chapter

I. INTRODUCTION. . . . . . . . . . . . . . . . . . . . . . . . 1

II. MARKOV CHAINS WITH COUNTABLY MANY STATES. . . . . . . . . . 4

Preliminaries: Basic Definitions, Notation, and Relationships
Asymptotic Behavior of the Transition Probabilities
Absolute Probabilities and Initial Distributions
Transient States
Theorems on the Classification of States

III. SOME ASPECTS OF QUEUEING THEORY . . . . . . . . . . . . . . 104

Classification of Queueing Systems
The Imbedded Markov Chain
The Random Arrival Queueing System
The Random Departure Queueing System

APPENDIX I . . . . . . . . . . . . . . . . . . . . . . . . . . 131

APPENDIX II. . . . . . . . . . . . . . . . . . . . . . . . . . 134

LITERATURE CITED . . . . . . . . . . . . . . . . . . . . . . . 138

# CHAPTER I

## INTRODUCTION

The purpose of this paper is to give a self-contained account of the basic theory of Markov chains having a countable number of states and to show how these Markov chains play a role in queueing theory.

The Markov chain is a special case of the Markov process, introduced in 1907 by A. A. Markov using a discrete parameter and a finite number of states. In 1936 the countable case was introduced by Kolmogorov, whose work was followed closely by that of Doeblin whose contributions pervade all parts of the Markov theory. The results of Chapter II are known in the literature but there appears to be no single reference which contains all of these results, or takes full advantage of the more precise classifications made possible by combining the results of various approaches. The objective of Chapter II is to combine the work of Chung [1], Feller [2] and Loève [1], the primary references for the chapter, in order to obtain stronger results where possible, as well as more detailed descriptions, and to give modified and more detailed proofs with the goal of greater clarity. The terminology follows Feller [2] in cases of conflict, since, unfortunately, there is no standard terminology in the literature.

In Chapter II we investigate a sequence $\{x_n\}$ of random variables which satisfies the Markov assumption: For every finite sequence $n_1 < n_2 < \ldots < n_r < n$ of indices and states $E_{j_{n_1}}, \ldots, E_{j_{n_r}}, E_k,$

$$P\{x_n = E_k | x_{n_1} = E_{j_{n_1}}, \ldots, x_{n_r} = E_{j_{n_r}}\} = P\{x_n = E_k | x_{n_r} = E_{j_{n_r}}\},$$

whenever the conditional probability on the left side is defined. Only Markov chains with stationary transition probabilities are considered. Various definitions, notation, preliminary relationships and a comparison of terminology are presented and then attention is focused on three major issues: (1) the asymptotic behavior of the n-step transition probabilities $p_{ij}^{(n)}$, (2) conditions for existence of a stationary probability distribution, and (3) theorems which classify the states of a Markov chain according to various criteria. Results are interpreted in terms of flow from state to state and periodicity is treated in the general case. Following Chung [1], Doob [1], Kolmogorov [1], and Loève [1], the proof of existence of such Markov chains is sketched in Appendix I. In Appendix II a lemma in elementary number theory, which does not appear to occur in the basic references in number theory, is established.

In Chapter III the basic concepts of queueing theory are introduced with a brief discussion of queueing systems in general in order that the results of Chapter II may be applied to two special queueing systems. (The first major contributions to the theory of queues were made by A. K. Erlang in 1908.) Queueing systems are classified according to the arrival pattern, service mechanism, and queue discipline. Following primarily Kendall [2], [3], and Foster [1], the method of the imbedded Markov chain is presented. In general this method does not lead to a Markov chain with a countable number of states, but for the queueing system with random arrivals, unspecified (but identical and independent) service time distributions, and one server, and the queueing system with an unspecified arrival distribution (but identical and independent inter-arrival time distributions), random service times, and a finite number of servers, the method does lead to such a Markov chain. It is shown, using the above references, that the imbedded Markov chain for each of the above queueing systems is positive, persistent null, or transient according as the ratio of average input to average output of the system is less than, equal to, or greater than one, and these results for the imbedded Markov chain are interpreted in terms of queue length.

# CHAPTER II

## MARKOV CHAINS WITH COUNTABLY MANY STATES

In this chapter we shall investigate Markov chains with countably many states. Thus we shall be concerned with a sequence $\{x_n\}$ of random variables which satisfies the Markov assumption: For every finite sequence $n_1 < n_2 < \ldots < n_r < n$ of indices, and states $E_{j_{n_1}}, \ldots, E_{j_{n_r}}$, $E_k$,

$$P\{x_n = E_k | x_{n_1} = E_{j_{n_1}}, \ldots, x_{n_r} = E_{j_{n_r}}\} = P\{x_n = E_k | x_{n_r} = E_{j_{n_r}}\} ,$$

whenever the conditional probability on the left side is defined. The Markov assumption states, in effect, that knowledge of the value of some one of the random variables $x_i$ renders subsequent random variables independent of random variables preceding $x_i$. Attention will be restricted to Markov chains with stationary transition probabilities, and in Appendix I we include a sketch of the proof of existence of such Markov chains.

After presenting various basic definitions, notation, and some preliminary relationships, we shall concentrate attention on three major issues: (1) the asymptotic behavior of the n-step transition probabilities $p_{ij}^{(n)}$, (2) conditions for existence of a stationary probability distribution, and (3) theorems which classify the states of a Markov chain according to various criteria. In particular, some of the theorems on classification of states will be applicable to the classification of states in various special queueing problems. Throughout the chapter results will be interpreted in terms of flow from state to state and periodicity will be treated in the general case. A lemma in elementary number theory used in some of the proofs is established in Appendix II. While this lemma is often mentioned, a proof does not appear to occur in the basic references in number theory.

The aim in this chapter is to give a self-contained account of the theory along the lines outlined above. While the results included here are known in the literature, a unified account with some modifications and improvements of theorems and proofs, and much more detailed arguments in many instances, seems to be desirable and worthwhile.

## Preliminaries: Basic Definitions, Notation, and Relationships

Let a countable set I with members $E_j$, $j = 0, 1, 2, \ldots$, be given, and define $\Omega$ to be the set of all infinite sequences $\omega = E_{j_o}, E_{j_1}, \ldots$ (the denumerable Cartesian product IXIX...). A cylinder set $S(E_{j_o},\ldots,E_{j_n})$ is the set of points $\omega\varepsilon\Omega$ with first $n + 1$ coordinates specified to be $E_{j_o}, \ldots, E_{j_n}$, respectively.

We shall be concerned with a probability space $(\Omega,\zeta,P)$, where $\zeta$ is the smallest Borel field containing the cylinder sets $S(E_{j_o},\ldots,E_{j_n})$ for all choices of $n \geq 0$ and all choices of $E_{j_o},\ldots,E_{j_n}$, and P is a probability measure defined on $\zeta$. The nth coordinate $E_{j_n}$ of $\omega$ will be referred to as the "state at the nth step," and the states will be labeled so that each $E_{j_n}$ is a non-negative integer. (Note: The coordinates are numbered 0,1,... .)

Consider the sequence of random variables $\{x_n\}$, $n = 0,1,\ldots$, with $x_n(\omega) = E_{j_n}$, the nth coordinate of $\omega$. The random variables are said to form a *Markov chain* (with a countable number of states, since the set I of states is countable) if the following condition is satisfied:

Markov Assumption: The conditional probability that $x_n = E_{j_n}$, given the values of a finite set of the preceding random variables, depends only on the value of the last given random variable. In symbols, for every finite set of indices $n_1 < \ldots < n_r < n$ and states $E_{j_{n_1}}, \ldots, E_{j_{n_r}}$, $E_k$, we have

$$P\{x_n = E_k | x_{n_1} = E_{j_{n_1}}, \ldots, x_{n_r} = E_{j_{n_r}}\} = P\{x_n = E_k | x_{n_r} = E_{j_{n_r}}\},$$

whenever the conditional probability on the left side is defined. (The argument of $x_n$ will generally be omitted; thus, the assertion $x_n = E_k$ means that the value of the function $x_n$ at the point $\omega$ is $E_k$.)

The Markov assumption clearly does not require that the random variables be independent. Thus the state $E_k$ is not associated with a fixed probability $p_k$ at an arbitrary step $n \geq 0$, but rather with the probability $p_k = P\{x_o = E_k\}$, and probabilities $p_{jk}^{m,n-m}$ for $n \geq 1$, where $p_{jk}^{m,n-m}$ is the probability of passage from state $E_j$ at the mth step to state $E_k$ at the nth step, $n > m$, that is, passage in n-m steps. (Note: Superscripts are not to be confused with exponents.) The number $p_{jk}^{m,n-m}$ is the P-measure of the set of $\omega$ for which $x_m(\omega) = E_j$ and

$x_n(\omega) = E_k$, divided by the P-measure of the set of $\omega$ for which $x_m(\omega) = E_j$, if this second number is positive, and is undefined otherwise. This is, of course, what is meant by $P\{x_n = E_k | x_m = E_j\}$.

Probabilities $p_{jk}^{m,n-m}$ which depend on the difference n-m, the number of steps from state $E_j$ to state $E_k$, but *not* on the individual values of m and n are called *stationary transition probabilities*. These probabilities may be written as $p_{jk}^{(n-m)}$, so that $p_{jk}^{(n-m)}$ denotes the probability of passage from state $E_j$ to state $E_k$ in n-m steps. In this paper we shall consider only stationary transition probabilities.

We may assume that if $E_i \varepsilon I$ then $P\{x_n = E_i\} > 0$ for some $n \geq 0$, since otherwise the set of all $\omega$ with $E_i$ appearing as a coordinate would be a countable union of sets of P-measure zero and thus would have P-measure zero, and consequently $E_i$ would be of no interest. Then $p_{ij}^{(n)}$ is defined for every $E_i, E_j \varepsilon I$ and $n \geq 1$, since for some $m \geq 0$ it is true that $P\{x_m = E_i\} > 0$, in which case

$$p_{ij}^{(n)} = P\{x_{m+n} = E_j | x_m = E_i\} = \frac{P\{x_{m+n} = E_j,\ x_m = E_i\}}{P\{x_m = E_i\}} .$$

If $P\{x_r = E_i\} = 0$ then we cannot determine $P\{x_{n+r} = E_j | x_r = E_i\}$ from

$$P\{x_{n+r} = E_j, x_r = E_i\} = P\{x_{n+r} = E_j | x_r = E_i\}\ P\{x_r = E_i\}$$

by division; but, since we are assuming stationary transition probabilities,

$$P\{x_{n+r} = E_j | x_r = E_i\} = p_{ij}^{(n)} .$$

It will be convenient later to define

$$p_{ij}^{(0)} = \delta_{ij} ,$$

the Kronecker delta. The probability $p_{ij}^{(1)}$ will be denoted simply by $p_{ij}$ .

Since $p_{ij}^{(n)}$ is a probability, and the probability of passage from $E_i$ to some state in n steps is one we have

$$p_{ij}^{(n)} \geq 0$$

and

$$\sum_j p_{ij}^{(n)} = 1 ,$$

where an unspecified summation should be understood to extend over all states.

The finite-dimensional joint probabilities

$$P\{x_o = E_{i_o} , \ldots, x_n = E_{i_n}\}$$

for all $n \geq 0$ and all $E_{i_r} \varepsilon I$, (i.e., the probability measures of the cylinder sets) can be expressed as a product of conditional probabilities as follows:

$$P\{x_o = E_{i_o}\}P\{x_1 = E_{i_1}|x_o = E_{i_o}\}P\{x_2 = E_{i_2}|x_o = E_{i_o}, x_1 = E_{i_1}\}$$

$$\ldots P\{x_n = E_{i_n}|x_o = E_{i_o}, \ldots, x_{n-1} = E_{i_{n-1}}\}$$

by the definition of conditional probability, whether or not the sequence of random variables forms a Markov chain. We define this product to be zero if any of the conditional probabilities are undefined, in agreement with the original joint probability.

If the sequence of random variables forms a Markov chain, the above expression reduces, by the Markov assumption, to

$$P\{x_o = E_{i_o}\}P\{x_1 = E_{i_1}|x_o = E_{i_o}\}P\{x_2 = E_{i_2}|x_1 = E_{i_1}\}$$

$$\ldots P\{x_n = E_{i_n}|x_{n-1} = E_{i_{n-1}}\} = p_{i_o} p_{i_o i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n} .$$

The matrix $P = [p_{ij}]$ is called the *transition probability matrix* of the Markov chain and $\{p_i, E_i \varepsilon I\}$ the *initial distribution*, with each $p_i \geq 0$ and $\sum_i p_i = 1$ . The matrix P is called a *stochastic matrix*; it has the characteristic properties:

$$p_{ij} \geq 0 , \qquad \sum_j p_{ij} = 1$$

for every i and j. (If in addition $\sum_i p_{ij} = 1$ for every j, then P is said to be *doubly stochastic.*)

That a probability measure P can be defined on $\zeta$ so that the given sequence of random variables satisfies the Markov assumption and so that P agrees with prescribed probabilities $\{p_i\}$ and $[p_{ij}]$ for the initial distribution and transitions, is not obvious. As is indicated in Chung [1] and Doob [1], using Kolmogorov's extension theorem (Kolmogorov [1], Loève [1]), we see that for the $\Omega$ and $\zeta$ given above, such a P always exists. These results are sketched in Appendix I. Throughout the remainder of this paper the probability space $(\Omega,\zeta,P)$ and state space I will be understood.

If there are only finitely many states we have a finite-dimensional matrix P and a *finite-dimensional Markov chain* with stationary transition probabilities. Much simplification is possible in the finite case. (See Kemeny and Snell [1], and Fréchet [1], in this connection.)

We shall now establish a basic result, and proceed to classify the states of a Markov chain.

Theorem 1 (Chapman-Kolmogorov Equations)

$$p_{ij}^{(n+m)} = \sum_k p_{ik}^{(m)} p_{kj}^{(n)} \qquad (m=0,1,\ldots,\ n=0,1,\ldots\ .)$$

or, in matrix notation,

$$P^{m+n} = P^m P^n \ .$$

Proof.

$$p_{ij}^{(n+1)} = P\{x_{r+n+1} = E_j | x_r = E_i\}$$

$$= \sum_k P\{x_{r+1} = E_k | x_r = E_i\} P\{x_{r+n+1} = E_j | x_{r+1} = E_k\}$$

$$= \sum_k p_{ik} p_{kj}^{(n)}, \qquad n = 0,1,\ldots .$$

We now use induction on m. The case m = 0 is trivial and for m = 1 the theorem is true by earlier remarks. Assume that the theorem holds for m = s. Then

$$p_{ij}^{(n+(s+1))} = p_{ij}^{(1+(s+n))}$$

$$= \sum_k p_{ik} p_{kj}^{(s+n)}$$

$$= \sum_k p_{ik} (\sum_u p_{ku}^{(s)} p_{uj}^{(n)})$$

$$= \sum_u p_{uj}^{(n)} (\sum_k p_{ik} p_{ku}^{(s)})$$

$$= \sum_u p_{uj}^{(n)} \ p_{iu}^{(s+1)}$$

using a standard theorem on interchange of order of summation (for example, Apostol [1], Theorem 12-43). Thus the result holds for $m = s+1$, and the induction is complete.

Definition 1. The state $E_j$ is a *consequent* of the state $E_i$ if $p_{ij}^{(n)} > 0$ for some $n \geq 1$, and we write $E_i \to E_j$ to indicate that $E_j$ is a consequent of $E_i$.

Definition 2. If, whenever $E_j$ is a consequent of $E_i$, we also have that $E_i$ is a consequent of $E_j$, then $E_i$ is called an *essential* state. That is, $E_i \to E_j$ implies $E_j \to E_i$. Otherwise, $E_i$ is called *inessential*, in which case it is possible to leave $E_i$ and have a zero probability of ever returning. Thus if $E_i$ is inessential there exists an $E_j$ such that $E_i \to E_j$ but $E_j \not\to E_i$.

Definition 3. States $E_i$ and $E_j$ *communicate* if $E_i \to E_j$ and $E_j \to E_i$. For communicating states $E_i$ and $E_j$ we shall write $E_i \approx E_j$.

Theorem 2. The following relations hold:

(i) $E_i \approx E_j$ implies $E_j \approx E_i$.

(ii) $E_i \approx E_j$ and $E_j \approx E_k$ implies $E_i \approx E_k$.

(iii) If $E_i \approx E_j$ for some $j$ then $E_i \approx E_i$.

Proof. (i) is immediate from the definition of $\approx$. For (ii) note that for some m, n, r, s we know $p_{ij}^{(n)} p_{ji}^{(m)} p_{jk}^{(r)} p_{kj}^{(s)} > 0$, where $m, n, r, s \geq 1$. Thus, by Theorem 1,

$$p_{ik}^{(n+r)} = \sum_u p_{iu}^{(n)} p_{uk}^{(r)} \geq p_{ij}^{(n)} p_{jk}^{(r)} > 0 \quad \text{so} \quad E_i \to E_k ,$$

and

$$p_{ki}^{(m+s)} = \sum_u p_{ku}^{(s)} p_{ui}^{(m)} \geq p_{kj}^{(s)} p_{ji}^{(m)} > 0 \qquad \text{so} \qquad E_k \to E_i \; .$$

If $E_i \approx E_j$ for some j then $E_j \approx E_i$ by (i) and $E_i \approx E_i$ by (ii). Since it is not necessarily true that $E_i \approx E_i$, $\approx$ is not an equivalence relation. However, by assertion (iii) of Theorem 2, this occurs only when $E_i$ communicates with no other state, so by treating these noncommunicating states separately we may partition the remaining states into equivalence classes by the following definition, and say that $E_i$ and $E_j$ are equivalent if $E_i \approx E_j$.

Definition 4. If $E_i$ does not communicate with any state it forms a *class* $C_i$ by itself. If $E_i$ communicates with at least one state (possibly itself), all the states with which it communicates form a *class* $C_i$.

Theorem 3. $E_i \varepsilon C_i$; two classes are either identical or disjoint.

Proof. By definition of $C_i$ and by (iii) of Theorem 2, $E_i \varepsilon C_i$. Assume $C_i$ and $C_j$ are not disjoint, $i \neq j$. If $E_i$, say, does not communicate with any state then the classes are obviously disjoint, so $E_i$ and $E_j$ belong to classes of the second type in Definition 4. Thus, if $E_k \varepsilon C_i$ and $E_k \varepsilon C_j$ we know $E_i \approx E_k$, $E_k \approx E_j$ so by (ii) of Theorem 2, $E_i \approx E_j$.

If $E_r \varepsilon C_i$, then $E_r \approx E_i \approx E_j$ so $E_r \varepsilon C_j$.

If $E_r \varepsilon C_j$, then $E_r \approx E_j \approx E_i$ so $E_r \varepsilon C_i$.

Thus, $C_i$ and $C_j$ are identical whenever they have an element in common.

<u>Theorem 4</u>. If $E_i$ is essential and $E_j$ is inessential, then $p_{ij}^{(n)} = 0$ for all n.

<u>Proof</u>. Suppose $p_{ij}^{(n)} > 0$ for some n. The state $E_j$ is inessential so there is an $E_k$ such that $p_{jk}^{(m)} > 0$ and $p_{kj}^{(u)} = 0$ for every u. By Theorem 1,

$$p_{ik}^{(n+m)} = \sum_r p_{ir}^{(n)} p_{rk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0 \; .$$

$E_i$ is essential, so $p_{ki}^{(s)} > 0$ for some s. But then

$$p_{kj}^{(n+s)} = \sum_r p_{kr}^{(s)} p_{rj}^{(n)} \geq p_{ki}^{(s)} p_{ij}^{(n)} > 0 \; ,$$

which is a contradiction. Thus, $p_{ij}^{(n)} = 0$ for all n.

<u>Theorem 5</u>. Either all the states in a particular class are essential or they are all inessential.

<u>Proof</u>. This follows from Theorem 4 since an inessential state and an essential state cannot communicate.

On account of Theorem 5, we may call a class essential or inessential according as all the states in it are essential or inessential.

In an inessential class we may pass from any state in the class to any other state in the class (by the definition of class), and we may leave the class, never to return (by the definition of inessential). Once we enter an essential class we must remain in that class forever, since to pass to an inessential class is forbidden by Theorem 4, and to pass to another essential class is impossible by the disjointness assertion of Theorem 3 and the definition of essential.

Definition 5. A set of states C is *closed* if no state outside C can be reached from any state $E_j$ in C.

Definition 6. A set of states C is *open* if from every state in C it is possible to reach a state outside C.

Thus, an essential class is closed, an inessential class is open.

Definition 7. The smallest closed set containing a set C is called the *closure* of C.

Definition 8. A single state $E_j$ forming a closed set will be called an *absorbing state*. (This is clearly the case if and only if $p_{jj} = 1$.)

Definition 9. A Markov chain is *irreducible* if there exists no closed set other than the set of all states.

We see that C is closed if and only if $p_{jk} = 0$ whenever $E_j \varepsilon C$ and $E_k \notin C$, for then, using Theorem 1, $p_{jk}^{(n)} = 0$ for every n.

If in the matrices $P^n$, n = 1,2,..., all rows and all columns corresponding to states outside the closed set C are deleted there remain stochastic matrices for which the fundamental relations of Theorem 1 again hold, that is, we have a Markov chain defined on C. We may thus study these subchains independently of all other states.

Theorem 6. A Markov chain is irreducible if and only if every state can be reached from every other state. Thus, the states of an irreducible chain form a single class.

Proof. If the chain is irreducible, and $E_k$ is inaccessible from $E_j$, delete $E_k$ and all those states leading to $E_k$ with positive probability. Then $E_j$ and the other remaining states form a closed set properly included in the set of all states, contrary to the assumption that the chain is irreducible. The converse is immediate.

Having considered the flow from state to state and class to class, we may now gain further insight by considering the probability of ultimately passing from state $E_i$ to state $E_j$.

We denote by $f_{ij}^{(n)}$ the probability that starting from $E_i$ we reach $E_j$ for the <u>first</u> time after n steps. In particular, $f_{jj}^{(n)}$ is the probability that the first return to $E_j$ occurs on the nth step. We define

$$f_{ij}^{(o)} = 0$$

for all i (including i = j), and notice that

$$f_{jj}^{(1)} = p_{jj}^{(1)} = p_{jj} .$$

In general, for n > 1, we have

$$p_{jj}^{(n)} = f_{jj}^{(n)} + f_{jj}^{(n-1)}p_{jj}^{(1)} + f_{jj}^{(n-2)}p_{jj}^{(2)} + \dots + f_{jj}^{(1)}p_{jj}^{(n-1)}$$

by the total probability formula for disjoint events. Hence,

$$f_{jj}^{(n)} = p_{jj}^{(n)} - \sum_{u=1}^{n-1} f_{jj}^{(u)}p_{jj}^{(n-u)} . \qquad (1)$$

Also,

$$f_{ij}^{(1)} = p_{ij}$$

and

$$f_{ij}^{(n)} = p_{ij}^{(n)} - \sum_{u=1}^{n-1} f_{ij}^{(u)} p_{jj}^{(n-u)} \qquad (2)$$

for $n > 1$.

Then

$$f_{jj} = \sum_{n=1}^{\infty} f_{jj}^{(n)}$$

is the probability of ultimate return to $E_j$ (starting from $E_j$), and

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

is the probability that starting from $E_i$ the state $E_j$ is ever reached.

Theorem 7. $E_i$ and $E_j$ communicate if and only if $f_{ij} > 0$ and $f_{ji} > 0$.

Proof.

$$\sup_{n\geq 1} \; p_{ij}^{(n)} \leq \sum_{n=1}^{\infty} f_{ij}^{(n)} = f_{ij}$$

and

$$\sum_{n=1}^{\infty} f_{ij}^{(n)} \le \sum_{n=1}^{\infty} p_{ij}^{(n)}$$

since

$$f_{ij}^{(n)} \le p_{ij}^{(n)}$$

for each n.

Thus

$$\sup_{n\ge 1} \quad p_{ij}^{(n)} \le f_{ij} \le \sum_{n=1}^{\infty} p_{ij}^{(n)} .$$

If $E_i$ and $E_j$ communicate, then $p_{ij}^{(n)} > 0$ for some $n \ge 1$, so $f_{ij} > 0$. Similarly, $f_{ji} > 0$.

If $f_{ij} > 0$ and $f_{ji} > 0$ , then

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} > 0 ,$$

so $p_{ij}^{(n)} > 0$ for some $n \ge 1$. A similar argument shows that $p_{ji}^{(m)} > 0$ for some $m \ge 1$, so $E_i$ and $E_j$ communicate.

The *passage time* $v_{jk}$ from state $E_j$ to state $E_k$ takes on values $m = 1,2,\ldots$ with probability $f_{jk}^{(m)}$. We define the *expected passage time*

$$\mu_{jk} = \sum_{n=1}^{\infty} nf_{jk}^{(n)} + \infty(1-f_{jk})$$

and the *mean recurrence time*

$$\mu_{jj} = \sum_{n=1}^{\infty} nf_{jj}^{(n)} + \infty(1-f_{jj}) \tag{3}$$

where $\infty \cdot c = 0$ or $\infty$ according as $c = 0$ or $c > 0$. The mean frequency of returns to $E_j$ is $1/\mu_{jj}$ where $1/\infty = 0$.

The states of a Markov chain may further be classified in terms of the above quantities.

Definition 10. $E_j$ is *transient* if there is a positive probability of never returning to the state $E_j$, i.e., if $f_{jj} < 1$. $E_j$ is *persistent* if there is probability one of returning to the state $E_j$, i.e., if $f_{jj} = 1$. In addition, $E_j$ is *null* or *positive* according as $\mu_{jj}$ is infinite or finite. (The term, *recurrent*, is often used in place of persistent.)

Automatically, transient states are null, since $\mu_{jj} = \infty$ if $1 - f_{jj} > 0$. A special type of transient state occurs when $f_{jj} = 0$. In this case we say that $E_j$ is a *noreturn* state since there is a zero probability of ever returning. If there is a positive probability of returning to state $E_j$, i.e., $f_{jj} > 0$, then we say that $E_j$ is a *return* state. Note that a persistent state is always a return state, while a transient state may be return or noreturn.

Theorem 8. $E_j$ is a return or a noreturn state according as $p_{jj}^{(n)} > 0$ for at least one $n \geq 1$ or $p_{jj}^{(n)} = 0$ for all $n \geq 1$.

Proof. From the proof of Theorem 7 we have

$$\sup_{n\geq 1} p_{jj}^{(n)} \leq f_{jj} \leq \sum_{n=1}^{\infty} p_{jj}^{(n)} .$$

If $p_{jj}^{(n)} > 0$ for some $n \geq 1$ then $f_{jj} > 0$, so that $E_j$ is a return state. If $p_{jj}^{(n)} = 0$ for all $n \geq 1$ then $f_{jj} = 0$, and $E_j$ is a noreturn state.

We also say that $E_j$ is an *everreturn* state if for every $E_k$ such that $f_{jk} > 0$ we have $f_{kj} > 0$. Note that an everreturn state is the same as an essential state. For if $f_{jk} > 0$, then $f_{jk}^{(n)} > 0$ for some $n \geq 1$. Thus, $p_{jk}^{(n)} > 0$ so if $E_j$ is essential there exists an $m \geq 1$ such that $p_{kj}^{(m)} > 0$. Then there exists an $m_1 \leq m$ such that $f_{kj}^{(m_1)} > 0$ so $f_{kj} > 0$ and $E_j$ is an everreturn state. Conversely, suppose $p_{jk}^{(n)} > 0$ for some $n \geq 1$. Then for some $n_1 \leq n$, $f_{jk}^{(m_1)} > 0$ so $f_{jk} > 0$. But then $f_{kj} > 0$ so for some $m \geq 1$, $f_{kj}^{(m)} > 0$ so $p_{kj}^{(m)} > 0$ and $E_j$ is an essential state.

At this point we shall consider periodicity of states. For example, a sequence of experiments may be to flip a coin and roll a die alternately. The possible states are $E_o = 1$, $E_1 = 2$, ..., $E_5 = 6$, $E_6 = 7$ (=Heads), $E_7 = 8$ (=Tails), and the possible values of the random variables $x_{2n}$, $n \geq 0$, are $E_6$ and $E_7$, and of $x_{2n+1}$, $n \geq 0$, are $E_o$, $E_1$, ..., $E_5$. Then return to state $E_7$, say, is impossible except, perhaps, in 2, 4, ... steps.

Definition 11. The state $E_j$ has *period* $t_j \geq 1$ if $p_{jj}^{(n)} = 0$ whenever $n$ is not divisible by $t_j$, and $t_j$ is the greatest integer with this property. (That is, a return to $E_j$ is impossible except, perhaps, in $t_j$, $2t_j$, $3t_j$, ... steps.)

There are other ways in which we could define periodicity. If $E_j \approx E_j$ (so that $p_{jj}^{(n)} > 0$ for some $n \geq 1$), let $\tau_j$ be the greatest common divisor of all positive integers n for which $p_{jj}^{(n)} > 0$. Note that $p_{jj}^{(n)} > 0$ only if $n = mt_j$ so $t_j$ is a divisor of n. Thus, $\tau_j \geq t_j$. Also, if n is not divisible by $\tau_j$ then $p_{jj}^{(n)} = 0$ so $t_j \geq \tau_j$. Hence, $t_j = \tau_j$.

By Theorem 1, if $p_{jj}^{(k)} > 0$ then $p_{jj}^{(nk)} > 0$ for $n = 1,2,\ldots,$ since

$$p_{jj}^{((n+1)k)} = p_{jj}^{(nk+k)} = \sum_u p_{ju}^{(nk)} p_{uj}^{(k)} \geq p_{ju}^{(nk)} p_{jj}^{(k)} > 0 \ .$$

If r is the smallest integer greater than zero such that $p_{jj}^{(r)} > 0$ then $t_j \leq r$, and strict inequality may hold. For example, if $p_{jj}^{(n)} = 0$ except for $p_{jj}^{(6)}$, $p_{jj}^{(9)}$, $p_{jj}^{(12)}$, ..., then $r = 6$ but $t_j = 3$.

If $E_j \not\approx E_j$ then $p_{jj}^{(n)} = 0$ for all $n \geq 1$ so that $f_{jj} = 0$, $E_j$ is a noreturn state, and we say arbitrarily that the period $t_j = \infty$. This convention is implied by Definition 11.

In case $t_j > 1$, we speak of the *periodic state* $E_j$ with *period* $t_j$. If $t_j = 1$, $t_j$ is still called the period, but we will refer to the state $E_j$ as *nonperiodic* or *aperiodic*. The period $t_j$ clearly conveys the idea that state $E_j$ is accessible for return only in numbers of steps which are integral multiples of the period. Note that some steps in which $E_j$ is inaccessible may possibly be listed with those steps in which $E_j$ is accessible, for example $n = 1 \cdot t_j = 3$ above.

We can equivalently let $t_j$ be the greatest common divisor of all integers n for which $f_{jj}^{(n)} > 0$. To deduce this, let $\bar{t}_j$ = g.c.d. $\{n\geq 1: f_{jj}^{(n)} > 0\}$, and recall that $t_j$ = g.c.d. $\{n\geq 1: p_{jj}^{(n)} > 0\}$. Since $p_{jj}^{(n)} \geq f_{jj}^{(n)}$,

$$\{n\geq 1: \; f_{jj}^{(n)} > 0\} \subset \{n\geq 1: \; p_{jj}^{(n)} > 0\} ,$$

and thus $\bar{t}_j \geq t_j$. If $\bar{t}_j = 1$, then $t_j = 1$ and $\bar{t}_j = t_j = 1$. Suppose that $\bar{t}_j > 1$. Note that $f_{jj}^{(k)} = 0$ whenever k is not a multiple of $\bar{t}_j$. Now $p_{jj}^{(r)} = 0$ for $r = 1,\ldots,\bar{t}_j-1$, since

$$p_{jj}^{(r)} = f_{jj}^{(r)} + \sum_{v=1}^{r-1} f_{jj}^{(v)} p_{jj}^{(r-v)} \quad \text{for } r > 1 .$$

Suppose that $p_{jj}^{(k\bar{t}_j+r)} = 0$ for all $k = 0,\ldots,m$ and all $r = 1,\ldots,\bar{t}_j-1$. Then

$$p_{jj}^{((m+1)\bar{t}_j+r)} = \sum_{v=1}^{m+1} f_{jj}^{(v\bar{t}_j)} p_{jj}^{((m+1-v)\bar{t}_j+r)} = 0$$

for $r = 1,\ldots,\bar{t}_j-1$, by the induction hypothesis. It follows that $p_{jj}^{(k)} = 0$ when k is not a multiple of $\bar{t}_j$, and thus

$$\{n\geq 1: \; p_{jj}^{(n)} > 0\} \subset \{m\bar{t}_j: \; m=1,2,\ldots\} .$$

Hence $t_j \geq \bar{t}_j$. In view of the inequality $t_j \leq \bar{t}_j$ obtained earlier, it follows that $t_j = \bar{t}_j$ . (The case of infinite periods, not treated above, follows immediately from Theorem 8.)

Definition 12. A state which is positive and aperiodic will be called *ergodic*.

The following theorems present some fundamental properties of periodic states.

Theorem 9. All of the states in one class have the same period.

Proof. Let $E_i$ and $E_j$ belong to the same class, with $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$. By (iii) of Theorem 2, $E_i \approx E_i$, $E_j \approx E_j$, so $t_i$ and $t_j$ are finite. If $p_{ii}^{(v)} > 0$ for some $v > 0$, then

$$p_{jj}^{(m+v+n)} \geq p_{ji}^{(m)} p_{ii}^{(v)} p_{ij}^{(n)} > 0$$

and $p_{ii}^{(2v)} > 0$, so that $p_{jj}^{(m+2v+n)} > 0$. Thus $t_jK = m + v + n$ and $t_jL = m + 2v + n$ where K and L are integers. Hence $t_jL = t_jK + v$ and $v = t_j(L-K)$. Hence $t_j$ is a divisor of v for every $v > 0$ such that $p_{ii}^{(v)} > 0$. Thus $t_j \leq t_i$. By symmetry $t_i \leq t_j$, and thus $t_i = t_j$ .

To prove the next theorem we shall use the following lemma, a proof of which is given in Appendix II. (Actually, a weaker form of the lemma, such as that in Kemeny and Snell [1], would suffice but the lemma is commonly referred to in the following form.)

Lemma. If $a_1, a_2, \ldots, a_n$ are n distinct positive integers with greatest common divisor one, then any integer $N > a_1a_2\ldots a_n$ can be represented in the form

$$N = x_1a_1 + x_2a_2 + \dots + x_na_n$$

where the $x_i$ are positive integers.

Theorem 10. If $E_j$ is a return state, $p_{jj}^{(mt_j)} > 0$ for all sufficiently large integers m, that is for all m greater than a certain fixed integer.

Proof. By Theorem 8, $p_{jj}^{(n)} > 0$ for some $n \geq 1$. Let $n_1 \geq 1$ be the smallest such n. If $t_j < n_1$, there exists an integer $n_2 > n_1$ such that $p_{jj}^{(n_2)} > 0$ and $n_2$ is not divisible by $n_1$, and thus the greatest common divisor r of $n_1$ and $n_2$ is less than or equal to $n_1 - 1$, and $t_j \leq r$. Continuing in this manner, the process must terminate since $t_j \geq 1$. Thus there exists a finite set of distinct positive integers $n_1, n_2, \dots, n_k$ such that $t_j$ is their greatest common divisor and $p_{jj}^{(n_i)} > 0$ for $i = 1,2,\dots,k$. Then the set of distinct positive integers $n_1/t_j, n_2/t_j, \dots, n_k/t_j$ has greatest common divisor one. By the previous lemma any integer $m > 1/t_j^k(n_1n_2\dots n_k)$ can be represented in the form $m = x_1\, n_1/t_j + \dots + x_k\, n_k/t_j$ where the $x_i$ are positive integers. Thus $mt_j = x_1n_1 + \dots + x_kn_k$, and

$$p_{jj}^{(mt_j)} = p_{jj}^{(x_1n_1+\dots+x_kn_k)} \geq p_{jj}^{(x_1n_1)} \cdots p_{jj}^{(x_kn_k)} > 0 ,$$

since $p_{jj}^{(x_in_i)} > 0$ , $i = 1,2,\dots,k$.

Theorem 11. If $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$ then $m + n$ is divisible by $t_i$ and $t_j$. If, in addition, $p_{ij}^{(n')} > 0$ then $n \equiv n'$ (mod $t_i$). (Note that $t_i = t_j$ by Theorem 9.)

Proof. It follows from the inequality

$$p_{ii}^{(m+n)} = \sum_u p_{iu}^{(n)} p_{ui}^{(m)} \geq p_{ij}^{(n)} p_{ji}^{(m)} > 0$$

that $m + n$ is divisible by $t_i$. Similarly, $m + n$ is divisible by $t_j$. If $p_{ij}^{(n')} > 0$, then $n' + m$ is also divisible by $t_i$. Hence $k_1 t_i = m + n$ and $k_2 t_i = m + n'$ where $k_1$ and $k_2$ are positive integers. Thus $(k_1 - k_2) \cdot t_i = n - n'$, and $n \equiv n' \pmod{t_i}$.

Thus, if $C_i$ is not the type of class consisting of a single no-return state, to every $E_j$ in $C_i$ there corresponds a unique least integer $r_j$ such that $p_{ij}^{(n)} > 0$ implies that $n \equiv r_j \pmod{t_i}$. If $i = j$, then $r_j = 0$.

If $p_{ij}^{(n)} > 0$, and $E_i$ is a return state then

$$p_{ij}^{(kt_i+n)} = \sum_u p_{iu}^{(kt_i)} p_{uj}^{(n)} \geq p_{ii}^{(kt_i)} p_{ij}^{(n)} > 0$$

for all sufficiently large integers $k$, by Theorem 10. Note that $kt_i + n \equiv n \equiv r_j \pmod{t_i}$. For all sufficiently large $m$ such that $m \equiv r_j \pmod{t_i}$ we have $p_{ij}^{(m)} > 0$.

All states $E_j \varepsilon C_i$ with the same $r_j = r$ form a subclass $C_i(r)$ of $C_i$.

Theorem 12. Let $E_i$ be an essential state. Then

$$\sum_{E_j \varepsilon C_i(r)} p_{ij}^{(n)} = 1 \tag{4}$$

if $n \equiv r \pmod{t_i}$. In terms of the random variables,

$$P\{x_{k+n} \varepsilon C_i(r) | x_k = E_i\} = 1$$

if $n \equiv r \pmod{t_i}$.

Proof. $C_i$ is an essential class by Theorem 5. If $E_j \notin C_i$, then $p_{ij}^{(n)} = 0$ for all n. Then for n fixed

$$\sum_{E_j \varepsilon C_i} p_{ij}^{(n)} = 1 \ . \tag{5}$$

The period $t_i$ is finite since $E_i \approx E_i$, and thus Equation 5 is valid in particular for $n \equiv r \pmod{t_i}$, where r is defined above. But if $E_j \notin C_i(r)$ then $p_{ij}^{(n)} = 0$ for $n \equiv r \pmod{t_i}$, and we have Equation 4 for $n \equiv r \pmod{t_i}$.

Thus, once we enter an essential class $C_i$ at state $E_i$ we remain in $C_i$ forever and move in a cycle,

$$C_i(0) \rightarrow C_i(1) \rightarrow \ldots \rightarrow C_i(t_i - 1) \rightarrow C_i(0) \ .$$

Once again considering the probabilities of ultimate return and passage, we denote by $f_{jk}^{n}$ the probability, starting at state $E_j$, of passing through state $E_k$ at least n times. This probability is given by

$$f_{jk}^{n} = \sum_{m=1}^{\infty} f_{jk}^{(m)} f_{kk}^{n-1} = f_{jk} f_{kk}^{n-1} \ .$$

In particular, $f_{jj}^n$ is the probability of returning to state $E_j$ at least n times and is given by

$$f_{jj}^n = f_{jj}\, f_{jj}^{\,n-1} = (f_{jj})^2\, f_{jj}^{\,n-2} = \ldots = (f_{jj})^n ,$$

where parentheses are used to indicate an exponent.

The probability of returning to $E_j$ infinitely often is

$$F_{jj} = \lim_{n\to\infty} f_{jj}^{\,n} = \lim_{n\to\infty} (f_{jj})^n ,$$

which is zero or one according as $f_{jj} < 1$ or $f_{jj} = 1$. Thus the probability of returning to $E_j$ infinitely often is zero if $E_j$ is transient and is one if $E_j$ is persistent. Note that

$$F_{ij} = \lim_{n\to\infty} f_{ij}^n = \lim_{n\to\infty} f_{ij}\, f_{jj}^{\,n-1} = f_{ij}\, F_{jj} , \tag{6}$$

where $F_{ij}$ is the probability, starting at state $E_i$, of passing through state $E_j$ infinitely often. Thus $F_{ij}$ equals $f_{ij}$ or zero according as state $E_j$ is persistent or transient.

<u>Theorem 13</u>. If $E_i$ is a persistent state and $E_j \varepsilon C_i$, then $F_{ij} = 1$.

<u>Proof</u>.

$$P\{x_n = E_i \text{ for at least one } n \geq N;\ x_n \neq E_j \text{ for all } n \geq M \mid x_0 = E_i\}$$

$$= P\{x_N = E_i;\ x_n \neq E_j \text{ for all } n \geq M | x_o = E_i\}$$

$$+ \sum_{m=N+1}^{\infty} P\{x_m = E_i;\ x_n \neq E_i \text{ for } N \leq n < m;\ x_n \neq E_j \text{ for all } n \geq M | x_o = E_i\}$$

$$= P\{x_N = E_i;\ x_n \neq E_j \text{ for } M \leq n \leq N,\ x_n \neq E_j \text{ for all } n > N | x_o = E_i\}$$

$$+ \sum_{m=N+1}^{\infty} P\{x_m = E_i;\ x_n \neq E_i \text{ for } N \leq n < m;\ x_n \neq E_j \text{ for } M \leq n \leq N,\ x_n \neq E_j \text{ for all } n > N | x_o = E_i\}$$

$$\leq P\{x_N = E_i;\ x_n \neq E_j \text{ for } M \leq n \leq N | x_o = E_i\}\ P\{x_n \neq E_j \text{ for all } n > N | x_N = E_i\} + \sum_{m=N+1}^{\infty} P\{x_m = E_i;\ x_n \neq E_i \text{ for } N \leq n < m;$$

$$x_n \neq E_j \text{ for } M \leq n \leq N | x_o = E_i\}\ P\{x_n \neq E_j \text{ for all } n > m | x_m = E_i\}$$

using the Markov assumption. The inequality arises from the latter term since it is no longer required that $x_n \neq E_j$ for $N + 1 \leq n \leq m$.

The latter probabilities reduce to

$$P\{x_N = E_i;\ x_n \neq E_j \text{ for } M \leq n \leq N | x_o = E_i\}\ (1 - f_{ij})$$

$$+ \sum_{m=N+1}^{\infty} P\{x_m = E_i;\ x_n \neq E_i \text{ for } N \le n < m;\ x_n \neq E_j \text{ for } M \le n \le N | x_0 = E_i\} \cdot$$

$$(1 - f_{ij})$$

$$= P\{x_n = E_i \text{ for at least one } n \ge N;\ x_n \neq E_j \text{ for}$$

$$M \le n \le N | x_0 = E_i\}\ (1 - f_{ij}) .$$

Letting $N \to \infty$, we have

$$P\{x_n = E_i \text{ for infinitely many indices } n;\ x_n \neq E_j \text{ for all } n \ge M | x_0 = E_i\}$$

$$\le P\{x_n = E_i \text{ for infinitely many indices } n;\ x_n \neq E_j \text{ for all}$$

$$n \ge M | x_0 = E_i\}\ (1 - f_{ij}) .$$

But $E_i$, $E_j \in C_i$ and hence, by Theorem 7, $f_{ij} > 0$ and $1 - f_{ij} < 1$. Thus $A \le Aa$ where $0 \le a < 1$. Thus $A = 0$, i.e.,

$$P\{x_n = E_i \text{ for infinitely many indices } n;\ x_n \neq E_j$$

$$\text{for all } n \ge M | x_0 = E_i\} = 0 .$$

Since $E_i$ is persistent,

$$1 = F_{ii}$$

$$= P\{x_n = E_i \text{ for infinitely many indices } n | x_o = E_i\}$$

$$= P\{x_n = E_i \text{ for infinitely many indices } n;$$

$$x_n \neq E_j \text{ for all } n \geq M | x_o = E_i\}$$

$$+ P\{x_n = E_i \text{ for infinitely many indices } n;$$

$$x_n = E_j \text{ for some } n \geq M | x_o = E_i\} .$$

The first of the last two terms is zero and thus the last term equals one. Letting $M \to \infty$ ,

$$P\{x_n = E_i \text{ for infinitely many indices } n; x_n = E_j$$

$$\text{for infinitely many indices } n | x_o = E_i\} = 1 .$$

But $F_{ij}$ is greater than or equal to this probability. It follows that $F_{ij} = 1$ .

We now apply the above results to prove the important

Theorem 14. Either all the states in one class are persistent or they are all transient. (Thus we may classify $E_i$ by classifying any member of $C_i$.)

Proof. Suppose state $E_i$ is persistent, and let $E_j$ be any other state in the same class. By Theorem 13, $F_{ij} = 1$. Now, $F_{ij} = f_{ij} F_{jj}$, and thus $F_{jj} = 1$. Thus $f_{jj} = 1$, and state $E_j$ is persistent. That is, if one state in the class is persistent then all states in the class are persistent.

A *class* may then be called *persistent* or *transient* according as all the states in it are persistent or transient. As was indicated earlier, the classes are also described as essential or inessential. The relationship between these classifications will be presented in Theorem 16 of the next section.

## Asymptotic Behavior of the Transition Probabilities

This section begins with a basic theorem and a proof of the theorem which involves use of generating functions. This method will be useful again in this chapter, and also in the following chapter.

Theorem 15. State $E_j$ is transient if and only if

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty ,$$

in which case automatically

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$$

for each state $E_i$ .

State $E_j$ is persistent if and only if

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty .$$

Proof. Consider the sequence $a_o, a_1, a_2, \ldots$, where $0 \le a_i \le 1$, and let

$$A(s) = a_o + a_1 s + a_2 s^2 + \ldots ,$$

a well-defined function for every value of s for which the series converges. Then A(s) is called the generating function of the sequence $a_o, a_1, a_2, \ldots$ . Comparison with the geometric series shows that A(s) converges at least for $-1 < s < 1$ .

If $a_o + a_1 + a_2 + \ldots$ diverges (to $+\infty$ since all terms are nonnegative), then for any positive number M there exists an integer $n_1$ such that $a_o + a_1 + \ldots + a_{n_1} > M$ . Fix $n_1$ . Then

$$\lim_{s \to 1^-} (a_o + a_1 s + \ldots + a_{n_1} s^{n_1}) = a_o + a_1 + \ldots + a_{n_1} .$$

There exists a number $\delta$, $0 < \delta < 1$, such that $|(a_o + \ldots + a_{n_1}) - (a_o + \ldots + a_{n_1} s^{n_1})| < (a_o + \ldots + a_{n_1}) - M$ for $1 - \delta < s < 1$. That is, $A(s) \ge a_o + a_1 s + \ldots + a_{n_1} s^{n_1} > M$ for $1 - \delta < s < 1$ . But $M > 0$ is arbitrary, and thus

$$\lim_{s \to 1^-} A(s) = \infty$$

in this case.

If $a_o + a_1 + a_2 + \ldots$ converges, then for $\varepsilon > 0$ there exists an integer N such that for $n > N$,

$$\sum_{k=n}^{\infty} a_k < \varepsilon \ .$$

Let

$$A_n(s) = a_o + a_1 s + \ldots + a_n s^n \ .$$

Then for $n \geq N$ and $-1 \leq s \leq 1$ ,

$$|A(s) - A_n(s)| = |\sum_{k=n+1}^{\infty} a_k s^k| \leq \sum_{k=n+1}^{\infty} a_k < \varepsilon \ .$$

The sequence $\{A_n\}$ converges uniformly to A on $[-1,1]$. Each function $A_n$ is continuous on $[-1,1]$. Since the convergence is uniform on $[-1,1]$, A is continuous on $[-1,1]$. In particular, then,

$$\lim_{s \to 1^-} A(s) = A(1) < \infty \ .$$

Thus, in both cases,

$$\lim_{s \to 1^-} A(s) = A(1) .$$

(Note that if $A(1) < \infty$, the above result is a special case of Abel's limit theorem.)

We now define

$$F_{jk}(s) = \sum_{n=0}^{\infty} f_{jk}^{(n)} s^n \tag{7}$$

and

$$P_{jk}(s) = \sum_{n=0}^{\infty} p_{jk}^{(n)} s^n , \tag{8}$$

with $F_{jk}$ defined at least on $[-1,1]$ since $F_{jk}(1) = f_{jk}$ and $P_{jk}$ defined at least on $(-1,1)$.

From the basic recursion formula (2),

$$p_{ij}^{(n)} = \sum_{u=0}^{n} f_{ij}^{(u)} p_{jj}^{(n-u)} \tag{9}$$

for $n \geq 1$. Multiplying each side of (9) by $s^n$, $-1 < s < 1$, and summing:

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} s^n = \sum_{n=1}^{\infty} \sum_{u=0}^{n} f_{ij}^{(u)} p_{jj}^{(n-u)} s^n .$$

Then,

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} s^n - p_{ij}^{(0)} = \sum_{n=0}^{\infty} \sum_{u=0}^{n} f_{ij}^{(u)} p_{jj}^{(n-u)} s^n$$

$$= \sum_{n=0}^{\infty} \sum_{u=0}^{n} (f_{ij}^{(u)} s^u)(p_{jj}^{(n-u)} s^{n-u}) ,$$

which is the Cauchy product of $F_{ij}(s)$ and $P_{ij}(s)$, both of which are absolutely convergent on $(-1,1)$. Hence on $(-1,1)$,

$$P_{ij}(s) - p_{ij}^{(0)} = F_{ij}(s) P_{jj}(s) . \tag{10}$$

In particular

$$P_{jj}(s) - 1 = F_{jj}(s) P_{jj}(s)$$

or

$$P_{jj}(s) = \frac{1}{1 - F_{jj}(s)} \tag{11}$$

on $(-1,1)$.

Using the results mentioned for the function A,

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} = P_{jj}(1) = \lim_{s \to 1^-} P_{jj}(s)$$

$$= \lim_{s \to 1^-} \frac{1}{1 - F_{jj}(s)}$$

$$= \frac{1}{1 - f_{jj}} \quad .$$

Thus,

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty \quad \text{or} \quad \sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$$

according as $f_{jj} < 1$ or $f_{jj} = 1$ .

Also,

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = P_{ij}(1) - p_{ij}^{(0)} = \lim_{s \to 1^-} P_{ij}(s) - p_{ij}^{(0)}$$

$$= \lim_{s \to 1^-} F_{ij}(s)\, P_{jj}(s)$$

$$= F_{ij}(1)\, P_{jj}(1)$$

$$= f_{ij}\, P_{jj}(1) \, .$$

Thus, if $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$ , then $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$ for each state $E_i$ .

Theorem 16. An inessential class is transient, but an essential class may be persistent or transient.

Proof. Let state $E_i$ belong to an inessential class. Then state $E_i$ is inessential and there exists a state $E_j$ such that $p_{ij}^{(m)} > 0$ for some $m \geq 1$ and $p_{ji}^{(n)} = 0$ for all n. Then $f_{ji} = 0$, and (by (6)) $F_{ji} = 0$. Now, for any fixed integer $m \geq 1$,

$$F_{ii} = \sum_{u} p_{iu}^{(m)} F_{ui} = \sum_{u \neq j} p_{iu}^{(m)} F_{ui} + p_{ij}^{(m)} F_{ji}$$

$$= \sum_{u \neq j} p_{iu}^{(m)} F_{ui} \leq \sum_{u \neq j} p_{iu}^{(m)} = 1 - p_{ij}^{(m)} < 1 .$$

But $F_{ii} = 0$ or $F_{ii} = 1$ always, so $F_{ii} = 0$, $f_{ii} < 1$, and state $E_i$ is transient. Thus an inessential class is transient.

We complete the proof by considering two examples. Consider the Markov chain with states $E_0$, $E_1$, $E_{-1}$, $E_2$, $E_{-2}$, ..., and transition probabilities

$$p_{i,i+1} = p, \quad p_{i,i-1} = q$$

for all i, where $0 < p, q < 1$ and $p + q = 1$. From any state we may reach any other state,

$$p_{i,i+n}^{(n)} = (p)^n > 0 ,$$

and

$$p_{i+n,i}^{(n)} = (q)^n > 0 .$$

Thus the set of all states forms an essential class. To continue the analysis, we follow the outline of Rosenblatt [1], page 42. For n odd, $n \geq 1$, $p_{00}^{(n)} = 0$, since for each step $E_i$ to $E_{i+1}$, $i \geq 0$ ($E_i$ to $E_{i-1}$, $i \leq 0$), there must be a step $E_{i+1}$ to $E_i$ ($E_{i-1}$ to $E_i$) in order to return to $E_o$, that is, an even number of steps to return to $E_o$. For n even, $n \geq 2$, to go from $E_o$ to $E_o$ in n steps we must have transitions $E_i$ to $E_{i+1}$ n/2 times and transitions of the form $E_i$ to $E_{i-1}$ n/2 times. The number of distinct ways we can arrange the n/2 transitions $E_i$ to $E_{i+1}$ out of the total of n transitions is given by

$$\binom{n}{n/2} = \frac{n(n-1)\ \ldots\ (n - n/2 + 1)}{(n/2)!} ,$$

and the probability of each of these distinct ways of passing from $E_o$ to $E_o$ in n steps is $p^{n/2}\ q^{n/2}$. Hence

$$p_{oo}^{(n)} = \binom{n}{n/2} p^{n/2}\ q^{n/2}$$

for n even, $n \geq 2$. Then, using the notation of the proof of Theorem 15,

$$P_{oo}(s) = \sum_{n=0}^{\infty} p_{oo}^{(n)}\ s^n = \sum_{n=0}^{\infty} \binom{2n}{n} p^n\ q^n\ s^{2n} \tag{12}$$

$$= \sum_{n=0}^{\infty} (-1)^n\ p^n\ q^n\ 2^{2n} \binom{-\ 1/2}{n} s^{2n}$$

$$= \sum_{n=0}^{\infty} \binom{-\ 1/2}{n} (-s^2)^n\ (4pq)^n ,$$

the binomial series expansion for

$$(1 - 4pq\, s^2)^{-1/2} ,$$

for $-1 < s < 1$ since $0 < 4pq \le 1$. (We define $\binom{a}{o} = 1$.) To verify that

$$\binom{2n}{n} = (-1)^n\, 2^{2n} \binom{-1/2}{n} , \qquad (13)$$

note that for $n = 0$,

$$\binom{0}{0} = 1 \quad \text{and} \quad (-1)^o\, 2^o \binom{-1/2}{0} = 1 ,$$

and for $n = 1$,

$$\binom{2}{1} = 2 \quad \text{and} \quad (-1)^1\, 2^2 \binom{-1/2}{1} = 2 .$$

Assume (13) holds for $n = k$. Then

$$\binom{2(k+1)}{k+1} = \frac{(2k+2)!}{(k+1)!(k+1)!} = \frac{(2k+2)(2k+1)(2k)!}{(k+1)^2\, k!\, k!}$$

$$= \frac{4(k+1/2)}{k+1} \binom{2k}{k} = \frac{4(k+1/2)}{k+1} (-1)^k\, 2^{2k} \binom{-1/2}{k}$$

$$= 2^{2(k+1)} (-1)^{k+1} \frac{(-k-1/2)(-1/2)\ldots(1/2-k)}{(k+1)\; k!}$$

$$= 2^{2(k+1)} (-1)^{k+1} \binom{-1/2}{k+1} ,$$

and (13) holds for n = k + 1, and the induction is completed. For the first example, let

$$p = q = 1/2 .$$

Then, using (12),

$$\sum_{n=0}^{\infty} p_{oo}^{(n)} = P_{oo}(1) = \lim_{s\to 1^-} P_{oo}(s) = \lim_{s\to 1^-} (1 - s^2)^{-1/2} ,$$

which is infinite. By Theorem 15, $E_o$ is a persistent state. Hence, we have an essential class which is persistent. For the second example, let

$$p = 2/3, \quad q = 1/3 .$$

Then, using (12) again,

$$\sum_{n=0}^{\infty} p_{oo}^{(n)} = \lim_{s\to 1^-} (1 - \frac{8}{9} s^2)^{-1/2} ,$$

which is finite. Thus $E_o$ is a transient state. Hence we have in this case an essential class which is transient.

Even though $f_{ii} < 1$ for every state in a transient class $C_i$, it is still possible that $f_{ij} = 1$ for some state $E_j$ in $C_i$. For example, consider the Markov chain with states $E_o$, $E_1$, $E_2$, ..., and transition probabilities

$$p_{02} = 1,\ p_{11} = 1,\ p_{23} = p_{20} = p_{21} = 1/3\ ,$$

and

$$p_{i,i+1} = p_{i,i-1} = 1/2$$

for $i \geq 3$. Then

$$p_{01}^{(2)} = 1/3 > 0$$

but

$$p_{10}^{(n)} = 0$$

for all n. Thus $E_o$ is inessential, and hence transient by Theorem 16. From any state except $E_1$ we may reach any other state, so that the set of all states except $E_1$ forms a transient class, and

$$f_{02} = \sum_{n=1}^{\infty} f_{02}^{(n)} = f_{02}^{(1)} = 1\ .$$

The proof of the following theorem follows Loève [1], with numerous details inserted.

<u>Theorem 17</u>. A necessary and sufficient condition for state $E_j$ to be null is that

$$\lim_{n\to\infty} p_{jj}^{(n)} = 0 .$$

If $E_j$ is a positive state, then

$$p_{jj}^{(nt_j)} \to t_j/\mu_{jj} > 0$$

as $n \to \infty$ , and

$$p_{jj}^{(n)} = 0$$

for $n \not\equiv 0 \pmod{t_j}$ .

<u>Proof</u>. By definition of the period $t_j$, $p_{jj}^{(n)} = 0$ for all $n \not\equiv 0$ (mod $t_j$), and if and only if $E_j$ is null we have $\mu_{jj} = \infty$, and hence $t_j/\mu_{jj} = 0$, so it is only necessary to prove that

$$p_{jj}^{(nt_j)} \to t_j/\mu_{jj}$$

as $n \to \infty$. (The case $t_j = \infty$ is trivial.) We will break the proof into four parts.

(i) <u>Consider just the case $t_j = 1$</u>. Let

$$\alpha = \lim_n \sup p_{jj}^{(n)} ,$$

and note that $0 \le \alpha \le 1$. There exists a subsequence $(n_i)$ such that

$$p_{jj}^{(n_i)} \to \alpha .$$

Suppose now that p is an integer such that $f_{jj}^{(p)} > 0$. Corresponding to $\varepsilon > 0$ there is an integer $N > p$ for which

$$\sum_{m=N+1}^{\infty} f_{jj}^{(m)} < \varepsilon ,$$

since

$$\sum_{m=1}^{\infty} f_{jj}^{(m)} = f_{jj} \le 1 .$$

If $n_i > N$,

$$p_{jj}^{(n_i)} = \sum_{m=1}^{n_i} f_{jj}^{(m)} p_{jj}^{(n_i-m)} < f_{jj}^{(p)} p_{jj}^{(n_i-p)} + \sum_{\substack{m\le N \\ m\ne p}} f_{jj}^{(m)} p_{jj}^{(n_i-m)} + \varepsilon .$$

By the definition of $\alpha$ as an upper limit, there exists an integer $K \ge 1$ such that

$$p_{jj}^{(k)} < \alpha + \varepsilon \quad \text{for all } k \ge K .$$

If $n_i \geq K + N$, then

$$p_{jj}^{(n_i - m)} < \alpha + \varepsilon \quad \text{for} \quad m = 1, \ldots, N .$$

By the definition of the subsequence $(n_i)$, there exists an index $i_o$ such that

$$p_{jj}^{(n_i)} > \alpha - \varepsilon \quad \text{for all } i \geq i_o .$$

It follows that there is an index $i_1 \geq i_o$ such that

$$p_{jj}^{(n_i)} > \alpha - \varepsilon \quad \text{and} \quad p_{jj}^{(n_i - m)} < \alpha + \varepsilon \quad \text{for } m = 1, \ldots, N, \quad \text{if } i \geq i_1 .$$

If $i \geq i_1$, then

$$\alpha - \varepsilon < p_{jj}^{(n_i)} < f_{jj}^{(p)} \, p_{jj}^{(n_i - p)} + \sum_{\substack{1 \leq m \leq N \\ m \neq p}} (\alpha + \varepsilon) \, f_{jj}^{(m)} + \varepsilon ,$$

and hence

$$\alpha - \varepsilon < f_{jj}^{(p)} \, p_{jj}^{(n_i - p)} + (1 - f_{jj}^{(p)})(\alpha + \varepsilon) + \varepsilon .$$

It follows that

$$- \frac{3\varepsilon}{f_{jj}^{(p)}} + \alpha + \varepsilon < p_{jj}^{(n_i - p)} < \alpha + \varepsilon \quad (i \geq i_1) .$$

Thus

$$-\frac{3\varepsilon}{f_{jj}^{(p)}} + \alpha + \varepsilon \le \liminf_n p_{jj}^{(n_i-p)} \le \limsup_n p_{jj}^{(n_i-p)} \le \alpha + \varepsilon \ .$$

Since this assertion is true for every $\varepsilon > 0$, it follows that

$$\lim_{i\to\infty} p_{jj}^{(n_i-p)} = \alpha \ .$$

The same argument applies to show that

$$\lim_{i\to\infty} p_{jj}^{(n_i-2p)} = \lim_{i\to\infty} p_{jj}^{(n_i-p-p)} = \alpha \ ,$$

and, in general, that

$$\lim_{i\to\infty} p_{jj}^{(n_i-kp)} = \alpha \tag{14}$$

for each fixed positive integer k, assuming that $f_{jj}^{(p)} > 0$ .

(ii) Assume now that $f_{jj}^{(1)} > 0$. Then, by (14), for each fixed $k \ge 1$

$$p_{jj}^{(n_i-k)} \to \alpha \quad \text{as } i \to \infty \ .$$

If the state $E_j$ is transient, then $E_j$ is null. By Theorem 15,

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$$

and thus $p_{jj}^{(n)} \to 0$ as $n \to \infty$ . Consider now the case when $E_j$ is persistent.

Let

$$u_n = \sum_{m=n+1}^{\infty} f_{jj}^{(m)} \quad .$$

Note that $u_o = f_{jj} = 1$ (since $E_j$ is persistent), and $0 \leq u_n \leq 1$ for all n. Using (3), it follows that

$$\mu_{jj} = \sum_{m=1}^{\infty} m\, f_{jj}^{(m)} = \sum_{m=1}^{\infty} m(u_{m-1}-u_m) \quad .$$

There are now two possibilities. If $\mu_{jj} < \infty$ , then

$$n\, u_n = n \sum_{m=n+1}^{\infty} f_{jj}^{(m)} \leq \sum_{m=n+1}^{\infty} m\, f_{jj}^{(m)}$$

and hence $n\, u_n \to 0$ as $n \to \infty$ . Thus

$$\sum_{m=1}^{\infty} m(u_{m-1} - u_m) = \lim_{n\to\infty} \sum_{m=1}^{n} m(u_{m-1} - u_m)$$

$$= \lim_{n\to\infty}(u_o + u_1 + \ldots + u_{n-1} - n\, u_n) = \sum_{m=0}^{\infty} u_m \quad .$$

If $\mu_{jj} = \infty$ , the partial sums

$$\sum_{m=1}^{n} m(u_{m-1} - u_m) = u_o + \dots + u_{n-1} - n\, u_n$$

are unbounded above, and consequently $\sum_{m=0}^{\infty} u_m = \infty$ . Thus, in each case,

$$\mu_{jj} = \sum_{m=0}^{\infty} u_m \ . \tag{15}$$

Using the basic recursion formula (1),

$$p_{jj}^{(n)} = \sum_{m=1}^{n} f_{jj}^{(m)}\, p_{jj}^{(n-m)} = \sum_{m=1}^{n} (u_{m-1} - u_m)\, p_{jj}^{(n-m)}$$

Thus

$$\sum_{m=0}^{n} u_m\, p_{jj}^{(n-m)} = u_o\, p_{jj}^{(n)} + \sum_{m=1}^{n} u_{m-1}\, p_{jj}^{(n-m)} - p_{jj}^{(n)}$$

$$= \sum_{m=1}^{n} u_{m-1}\, p_{jj}^{(n-m)}$$

$$= \sum_{m=0}^{n-1} u_m\, p_{jj}^{(n-1-m)} \ .$$

If n = 1, $\sum_{m=0}^{1} u_m\, p_{jj}^{(n-m)} = u_o\, p_{jj}^{(0)} = 1$ .

If $\sum_{m=0}^{n} u_m \, p_{jj}^{(n-m)} = 1$ for some $n \geq 1$, then

$$\sum_{m=0}^{n+1} u_m \, p_{jj}^{(n+1-m)} = \sum_{m=0}^{n} u_m \, p_{jj}^{(n-m)} = 1 \quad .$$

It follows that

$$\sum_{m=0}^{n} u_m \, p_{jj}^{(n-m)} = 1 \qquad \text{for all } n \geq 1 \; . \tag{16}$$

If $n_i > n$, then

$$\sum_{m=0}^{n} u_m \, p_{jj}^{(n_i-m)} \leq \sum_{m=0}^{n_i} u_m \, p_{jj}^{(n_i-m)} = 1 \; .$$

But

$$p_{jj}^{(n_i-m)} \to \alpha$$

for each fixed $m \geq 0$ as $i \to \infty$ . Thus, for each first $n \geq 1$,

$$\sum_{m=0}^{n} u_m \, p_{jj}^{(n_i-m)} \to \alpha \sum_{m=0}^{n} u_m \leq 1 \qquad \text{as} \qquad i \to \infty \; .$$

If $\mu_{jj} = \sum_{m=0}^{\infty} u_m = \infty$ , then $\alpha = 0$ . If $\mu_{jj} < \infty$ , then $\alpha\mu_{jj} \leq 1$ .

In each case, then,

$$\alpha \le \frac{1}{\mu_{jj}}$$

(interpreting $1/\mu_{jj} = 0$ if $\mu_{jj} = \infty$) .

If

$$\beta = \liminf_{n} p_{jj}^{(n)} ,$$

a similar analysis shows that if $(n_i')$ is a subsequence such that

$$p_{jj}^{(n_i')} \to \beta ,$$

and if $f_{jj}^{(p)} > 0$ , then

$$p_{jj}^{(n_i'-kp)} \to \beta$$

for each fixed $k \ge 0$ and $i \to \infty$ . We assume that $f_{jj}^{(1)} > 0$ as before, and choose $p = 1$. If $\mu_{jj} = \infty$ , then $\alpha = 0$. In this case $\beta = 0$ also,

$$\alpha = \beta = \frac{1}{\mu_{jj}} .$$

Suppose now that $\mu_{jj} < \infty$ . Consider a fixed $n > 1$ and $n_i' > n$ .
Using (16), it follows that

$$1 = \sum_{m=0}^{n_i'} u_m \, p_{jj}^{(n_n'-m)} = \sum_{m=0}^{n} u_m \, p_{jj}^{(n_i'-m)} + \sum_{m=n+1}^{n_i'} u_m \, p_{jj}^{(n_i'-m)}$$

$$\leq \sum_{m=0}^{n} u_m \, p_{jj}^{(n_i'-m)} + \sum_{m=n+1}^{\infty} u_m \quad .$$

Letting $i \to \infty$ , it follows that

$$1 \leq \sum_{m=0}^{n} u_m \beta + \sum_{m=n+1}^{\infty} u_m \quad .$$

This inequality holds for each $n > 1$. Since $\mu_{jj} < \infty$ , the series $\sum u_m$ converges and

$$1 \leq \beta \, \mu_{jj} \; ,$$

so that

$$\beta \geq \frac{1}{\mu_{jj}} \quad .$$

Thus

$$\frac{1}{\mu_{jj}} \leq \beta \leq \alpha \leq \frac{1}{\mu_{jj}} \quad ,$$

which implies that

$$\alpha = \beta = \frac{1}{\mu_{jj}} \quad .$$

Thus $\lim_{n\to\infty} p_{jj}^{(n)}$ exists, and

$$\lim_{n\to\infty} p_{jj}^{(n)} = \frac{1}{\mu_{jj}} .$$

(iii) Now the assumption that $f_{jj}^{(1)} > 0$ must be removed. We are considering the case $t_j = 1$. As in the proof of Theorem 10, using the definition of period in the version preceding Definition 12, there exists a finite set of distinct positive integers $k_1, \ldots, k_r$ with greatest common divisor 1, such that $f_{jj}^{(k_i)} > 0$ for $i = 1, \ldots, r$. Using (14), for fixed positive integers $\alpha_1, \ldots, \alpha_r$ ,

$$p_{jj}^{(n_i-\alpha_1 k_1)} \to \alpha$$

$$p_{jj}^{(n_i-\alpha_1 k_1-\alpha_2 k_2)} \to \alpha ,$$

and, in general,

$$p_{jj}^{(n_i - \sum_{i=1}^{r} \alpha_i k_i)} \to \alpha .$$

By the number-theoretic lemma preceding Theorem 10, any integer $s > k_1 \ldots k_r = s_o$ can be written in the form

$$s = \sum_{i=1}^{r} \alpha_i k_i ,$$

where the coefficients $\alpha_i$ are positive integers. Thus, for each $s > s_o$,

$$p_{jj}^{(n_i - s)} \to \alpha \quad \text{as } i \to \infty .$$

Fix $n \geq 1$, and let $i_o$ be sufficiently large that $n_{i_o} > n + s_o + 1$, and consider $n_i$ with $i \geq i_o$. Using (16) once again,

$$\sum_{m=0}^{n} n_m \, p_{jj}^{(n_i - s_o - 1 - m)} \leq \sum_{m=0}^{n_i - s_o - 1} u_m \, p_{jj}^{(n_i - s_o - 1 - m)} = 1 .$$

Note that $s_o + 1 + m > s_o$ if $0 \leq m \leq n$. Hence, taking the limit as $i \to \infty$,

$$\sum_{m=0}^{n} u_m \alpha \leq 1 .$$

Arguing as in (ii), it follows that

$$\alpha \leq \frac{1}{\mu_{jj}} .$$

Similarly,

$$\beta \geq \frac{1}{\mu_{jj}} .$$

Thus

$$\alpha = \beta = \lim_{n\to\infty} p_{jj}^{(n)} = \frac{1}{\mu_{jj}} \quad ,$$

as in (ii).

(iv) <u>It remains to consider the case $t_j > 1$, the periodic case.</u> Suppose that $E_j$ is persistent. The transient case was discussed in (ii), and thus, by Theorem 16, it may be supposed that $E_j$ is in an essential class. By Theorem 12, transitions are in a cycle,

$$C_j(0) \to C_j(1) \to \ldots \to C_j(t_j - 1) \to C_j(0) \quad .$$

Now define a new Markov chain by letting one step correspond to $t_j$ steps in the original chain, that is,

$$p_{jj}^{'(n)} = p_{jj}^{(nt_j)} \quad .$$

Then

$$\sum_{n=1}^{\infty} f_{jj}^{'(n)} = \sum_{n=1}^{\infty} f_{jj}^{(nt_j)} = \sum_{n=1}^{\infty} f_{jj}^{(n)} = f_{jj} = 1 \ ,$$

Since $f_{jj}^{(k)} = 0$ for $k \neq nt_j$, and if $p_{jj}^{'(n)} > 0$ then $p_{jj}^{(nt_j)} > 0$, and thus state $E_j$ in the *new* chain is persistent and has period one. By the previous parts of this proof,

$$p_{jj}^{'(n)} \to \frac{1}{\mu_{jj}^{'}}$$

as $n \to \infty$ . But

$$\mu'_{jj} = \sum_{n=1}^{\infty} n\, f'_{jj}{}^{(n)} = \sum_{n=1}^{\infty} n\, f_{jj}{}^{(nt_j)} = \frac{1}{t_j} \sum_{n=1}^{\infty} (nt_j)\, f_{jj}{}^{(nt_j)}$$

$$= \frac{1}{t_j} \mu_{jj} ,$$

since $f_{jj}{}^{(n)} = 0$ for all other integers n. Thus

$$p_{jj}{}^{(nt_j)} = p'_{jj}{}^{(n)} \to \frac{1}{\mu'_{jj}} = \frac{t_j}{\mu_{jj}}$$

as $n \to \infty$ .

<u>Theorem 18</u>. For every state $E_j$:

(i) if state $E_k$ is null,

$$p_{jk}{}^{(n)} \to 0$$

as $n \to \infty$ ,

(ii) if state $E_k$ is positive,

$$p_{jk}{}^{(nt_k+r)} \to f_{jk}(r) \frac{t_k}{\mu_{kk}}$$

as $n \to \infty$ ,

(iii) and whatever be state $E_k$ ,

$$\frac{1}{n} \sum_{m=1}^{n} p_{jk}^{(m)} \to \frac{f_{jk}}{\mu_{kk}}$$

as $n \to \infty$ , where

$$f_{jk}(r) = \sum_{m=0}^{\infty} f_{jk}^{(mt_k+r)} ,$$

$r = 1, 2, \ldots, t_k$ . Here, $f_{jk}(r)$ is the probability of passage from $E_j$ to $E_k$ in $n \equiv r \pmod{t_k}$ steps, and

$$\sum_{r=1}^{t_k} f_{jk}(r) = f_{jk} .$$

(Note that if $t_k = 1$, in which case state $E_k$ is aperiodic, then $r = 1$, $f_{jk}(1) = f_{jk}$, and

$$p_{jk}^{(n)} \to \frac{f_{jk}}{\mu_{kk}}$$

as $n \to \infty$ .)

<u>Proof</u>. (i) If state $E_k$ is null, then by Theorem 17, $p_{kk}^{(n)} \to 0$ as $n \to \infty$ . From the basic recurrence relation (2),

$$p_{jk}^{(n)} = \sum_{u=1}^{n} f_{jk}^{(u)} p_{kk}^{(n-u)}$$

$$= \sum_{u=1}^{n'} f_{jk}^{(u)} p_{kk}^{(n-u)} + \sum_{u=n'+1}^{n} f_{jk}^{(u)} p_{kk}^{(n-u)}$$

$$\leq \sum_{u=1}^{n'} f_{jk}^{(u)} p_{kk}^{(n-u)} + \sum_{u=n'+1}^{n} f_{jk}^{(u)}$$

for $n' < n$. Thus,

$$\limsup p_{jk}^{(n)} \leq \sum_{u=n'+1}^{\infty} f_{jk}^{(u)}$$

for every $n'$. Since

$$\sum_{u=1}^{\infty} f_{jk}^{(u)} \leq 1 ,$$

it follows that

$$\limsup p_{jk}^{(n)} = 0 .$$

Hence

$$\lim_{n\to\infty} p_{jk}^{(n)} = 0 .$$

(ii) If state $E_k$ is positive, then

$$p_{kk}^{(nt_k)} \to \frac{t_k}{\mu_{kk}} > 0 ,$$

and

$$p_{kk}^{(nt_k+r)} = 0$$

for $1 \leq r < t_k$, by Theorem 17. Again using relation (2),

$$0 = p_{jk}^{(nt_k+r)} - \sum_{u=1}^{nt_k+r} f_{jk}^{(u)} p_{kk}^{(nt_k+r-u)}$$

$$= p_{jk}^{(nt_k+r)} - \sum_{u=0}^{n} f_{jk}^{(ut_k+r)} p_{kk}^{((n-u)t_k)} ,$$

since all other $p_{kk}^{(m)} = 0$ . Hence, for $n' < n$,

$$0 \leq p_{jk}^{(nt_k+r)} - \sum_{u=0}^{n'} f_{jk}^{(ut_k+r)} p_{kk}^{((n-u)t_k)}$$

$$= \sum_{u=n'+1}^{n} f_{jk}^{(ut_k+r)} p_{kk}^{((n-u)t_k)}$$

$$\leq \sum_{u=n'+1}^{n} f_{jk}^{(ut_k+r)} .$$

Hence,

$$\sum_{u=0}^{n'} f_{jk}^{(ut_k+r)} p_{kk}^{((n-u)t_k)} \leq p_{jk}^{(nt_k+r)} \leq \sum_{u=0}^{n'} f_{jk}^{(ut_k+r)} p_{kk}^{((n-u)t_k)}$$

$$+ \sum_{u=n'+1}^{n} f_{jk}^{(ut_k+r)} .$$

Letting $n \to \infty$, with $n'$ fixed,

$$\sum_{u=0}^{n'} f_{jk}^{(ut_k+r)} \frac{t_k}{\mu_{kk}} \le \liminf p_{jk}^{(nt_k+r)} \le \limsup p_{jk}^{(nt_k+r)}$$

$$\le \sum_{u=0}^{\infty} f_{jk}^{(ut_k+r)} \frac{t_k}{\mu_{kk}} \le \sum_{u=0}^{n'} f_{jk}^{(ut_k+r)} \frac{t_k}{\mu_{kk}} + \sum_{u=n'+1}^{\infty} f_{jk}^{(ut_k+r)} ,$$

for every $n'$. Thus

$$\frac{t_k}{\mu_{kk}} \sum_{u=0}^{\infty} f_{jk}^{(ut_k+r)} \le \liminf p_{jk}^{(nt_k+r)} \le \limsup p_{jk}^{(nt_k+r)}$$

$$\le \frac{t_k}{\mu_{kk}} \sum_{u=0}^{\infty} f_{jk}^{(ut_k+r)} .$$

Thus

$$p_{jk}^{(nt_k+r)} \to \frac{t_k}{\mu_{kk}} f_{jk}(r)$$

as $n \to \infty$ , noting the definition of $f_{jk}(r)$ .

(iii) The final part of the theorem is trivial if $t_k$ is infinite, since $p_{kk}^{(m)} = 0$ for all $m \ge 1$ implies, using (2), that $p_{jk}^{(m)} = f_{jk}^{(m)}$ . Hence

$$\sum_{m=1}^{n} p_{jk}^{(m)} = \sum_{m=1}^{n} f_{jk}^{(m)} \le f_{jk} \le 1$$

for all n. Thus,

$$\frac{1}{n} \sum_{m=1}^{n} p_{jk}^{(m)} \to 0$$

as $n \to \infty$ , and

$$\frac{f_{jk}}{\mu_{kk}} = 0$$

since $\mu_{kk} = \infty$ .

If $t_k$ is finite, let $n = (N + 1)t_k$, so that $n \to \infty$ if and only if $N \to \infty$ . Then

$$\frac{1}{n} \sum_{m=1}^{n} p_{jk}^{(m)} = \frac{1}{(N+1)t_k} \sum_{m=0}^{N} \sum_{r=1}^{t_k} p_{jk}^{(mt_k+r)} = \frac{1}{t_k} \sum_{r=1}^{t_k} \left( \sum_{m=0}^{N} \frac{p_{jk}^{(mt_k+r)}}{N + 1} \right) ,$$

and we know that

$$p_{jk}^{(mt_k+r)} \to \frac{t_k}{\mu_{kk}} f_{jk}(r)$$

as $m \to \infty$ , for each $r = 1, \ldots, t_k$, by parts (i) and (ii).

By a standard theorem concerning Cesàro summability (for example, Apostol [1], Theorem 12-48),

$$\sum_{m=0}^{N} \frac{p_{jk}^{(mt_k+r)}}{N + 1} \to \frac{t_k}{\mu_{kk}} f_{jk}(r)$$

as $N \to \infty$ . Thus,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{jk}^{(m)} = \frac{1}{t_k} \sum_{r=1}^{t_k} \frac{t_k}{\mu_{kk}} f_{jk}(r) = \frac{f_{jk}}{\mu_{kk}} .$$

Theorem 19. Either all states in one class are positive or they are all null.

Proof. If the class is transient, all states in the class are automatically null.

If the class is persistent, let $E_i$ and $E_j$ be two states in the class. For some n, m, $p_{ij}^{(n)} > 0$, $p_{ji}^{(m)} > 0$, and

$$p_{ii}^{(n+m+ut_j)} \geq p_{ij}^{(n)} \, p_{jj}^{(ut_j)} \, p_{ji}^{(m)} .$$

By Theorem 11, $n + m$ is divisible by $t_i$ $(=t_j)$, and by Theorem 17, if $E_j$ is positive then

$$p_{jj}^{(ut_j)} \to \frac{t_j}{\mu_{jj}} > 0$$

as $u \to \infty$ . Thus,

$$\lim_{u\to\infty} p_{ii}^{(n+m+ut_j)} \geq \lim_{u\to\infty} p_{ij}^{(n)} \, p_{jj}^{(ut_j)} \, p_{ji}^{(m)} > 0 .$$

But

$$\lim_{u\to\infty} p_{ii}^{(n+m+ut_j)} = \lim_{k\to\infty} p_{ii}^{(kt_j)} = \lim_{k\to\infty} p_{ii}^{(kt_i)} = \frac{t_i}{\mu_{ii}} ,$$

and thus

$$\frac{t_i}{\mu_{ii}} > 0 ,$$

and state $E_i$ is positive.

By Theorem 19, we may further denote a class as positive or null according as all states in the class are positive or null. From previous results (recall Theorem 16), we have the following relationships:

Inessential classes are transient, and transient classes are null. Positive classes are persistent, and persistent classes are essential. Actually, there are four distinct types of classes, namely inessential, essential-transient, persistent-null, and positive.

Since an irreducible Markov chain is a single class, by Theorem 6, the states are all transient, all persistent null, or all positive. In every case, by Theorem 9, all states have the same period, and every state can be reached from every other state.

Further, in any Markov chain, not necessarily irreducible, if $E_j$ is a persistent state and state $E_k$ is a consequent of $E_j$, then $E_k$ belongs to the same class as $E_j$ since we cannot leave an essential class (noting Theorem 16), and so $E_k$ is persistent. Also, $E_j$ and $E_k$ are both null or both positive and have the same period.

The set of all noreturn states forms individual classes since these states are noncommunicating, and the set of all return states forms individual classes by the remarks preceding Definition 4. It follows that in every Markov chain the persistent states can be divided in a unique manner into closed classes $C_1, C_2, \ldots$, such that from any state of a given class $C_u$ all states of that class, and no other states, can be reached. Each such $C_u$ forms an irreducible Markov chain by the remarks preceding Theorem 6. We may also have classes $T_1, T_2, \ldots$, which are not closed, containing inessential states from which states of $C_u$ can be reached. Any remaining states (or possibly the entire chain, as in the second example in the proof of Theorem 16) can be uniquely divided into closed classes $C_1', C_2', \ldots$, of transient, essential states, with the remarks above concerning $C_u$ also holding for $C_u'$. Each class $C_u$ (and $C_u'$) may be generated as the closure of any one of its states.

If the states of a closed class have period $t > 1$, the class splits into $t$ cyclic subclasses, $C(1), C(2), \ldots, C(t)$, such that passage from a state in $C(r)$ to a state in $C(r+1)$, where $C(t+1) = C(1)$, occurs with probability one, by Theorem 12. The states of $C(r)$ form an irreducible Markov chain (recall the last part of the proof of Theorem 17) whose transition matrix is obtained from

$$P^t = [p_{ij}^{(t)}]$$

by deleting all those elements $p_{jk}^{(t)}$ for which states $E_j$ or $E_k$ do not belong to $C(r)$.

The following theorem applies to the equivalence classes $C_u'$ above, and to those equivalence classes $C_u$ which are null.

Theorem 20. An essential null equivalence class $C$ is either empty or infinite.

Proof. Suppose C is non-empty. Then by Theorem 18,

$$p_{jk}^{(n)} \to 0$$

as $n \to \infty$ , for every state $E_k \varepsilon C$. But C is closed and thus

$$\sum_{E_k \varepsilon C} p_{jk}^{(n)} = 1$$

for $E_j \varepsilon C$ and for $n \geq 0$. Thus C is an infinite set.

Note that in a finite Markov chain there can be no essential null states, and thus the only possible states are inessential and positive. By Theorem 18,

$$p_{ij}^{(n)} \to 0$$

as $n \to \infty$ , if state $E_j$ is inessential, for each state $E_i$, but

$$\sum_j p_{ij}^{(n)} = 1$$

for each n. Thus at least one state in a *finite* Markov chain must be positive.

## Absolute Probabilities and Initial Distributions

Now consider the absolute probability of random variable $x_n$ having the value $E_j$, that is, of being in state $E_j$ on the nth step. Let

$$a_j = P\{x_o = E_j\} ,$$

where $a_j \geq 0$, for each state $E_j$, and

$$\sum_j a_j = 1 .$$

Recall from the first section of this chapter that

$$P\{x_o = E_{j_o} , x_1 = E_{j_1} , \ldots, x_n = E_{j_n} \} = a_{j_o} \, p_{j_o j_1} \cdots p_{j_{n-1} j_n} .$$

The probability of finding the system in state $E_k$ on the nth step is given by

$$a_k^{(n)} = \sum_j a_j \, p_{jk}^{(n)} ,$$

where $a_k^{(0)} = a_k$. We call the initial distribution $\{a_j\}$ *stationary* if the equations

$$a_j = \sum_i a_i \, p_{ij}$$

are satisfied. In this case, $a_k^{(1)} = a_k$, and if $a_k^{(n)} = a_k$ then

$$a_k^{(n+1)} = \sum_j a_j^{(n)} \, p_{jk} = \sum_j a_j p_{jk} = a_k .$$

Thus by induction, $a_k^{(n)} = a_k$ for each state $E_k$ and each $n \geq 0$ .

A natural question to ask is whether a stationary distribution exists for a given Markov chain, that is whether there is an initial distribution which remains invariant under transitions.

Theorem 21. An irreducible Markov chain satisfies one of the following two conditions:

(i) It is a null class and no stationary distribution exists.

(ii) It is a positive class and $\{a_k\} = \{\frac{1}{\mu_{kk}}\}$ gives the unique stationary distribution.

Proof. An irreducible chain forms a single class. By Theorem 19, the dichotomy (i), (ii) exists. We will now use some of the results on generating functions which were obtained in the course of the proof of Theorem 15. By (7), and a standard theorem on the differentiation of power series,

$$F'_{jj}(s) = \sum_{n=1}^{\infty} n\, f_{jj}^{(n)}\, s^{n-1}$$

for $-1 < s < 1$.

Using (8), and Theorem 15, if state $E_j$ is transient,

$$\lim_{s \to 1^-} P_{jj}(s) = P_{jj}(1) < \infty ,$$

and thus

$$\lim_{s \to 1^-} (1 - s)\, P_{jj}(s) = 0 .$$

If state $E_j$ is persistent, $f_{jj} = 1$ and

$$\mu_{jj} = \sum_{n=1}^{\infty} n\, f_{jj}^{(n)} .$$

If $\mu_{jj}$ is finite, or infinite,

$$\lim_{s \to 1^-} F'_{jj}(s) = F'_{jj}(1)^- = \mu_{jj} ,$$

just as was shown for the junction A in the proof of Theorem 15. But, since $F_{jj}(s)$ is continuous on the closed interval [-1,1], by the definition of the left-hand derivative,

$$F'_{jj}(1)^- = \lim_{s \to 1^-} \frac{F_{jj}(1) - F_{jj}(s)}{1 - s}$$

$$= \lim_{s \to 1^-} \frac{f_{jj} - F_{jj}(s)}{1 - s}$$

$$= \lim_{s \to 1^-} \frac{1 - F_{jj}(s)}{1 - s}$$

$$= \lim_{s \to 1^-} \frac{1}{(1 - s)\, P_{jj}(s)} ,$$

using (11).

In either case, transient or persistent,

$$\lim_{s \to 1^-} (1 - s) P_{jj}(s) = \frac{1}{\mu_{jj}} . \tag{17}$$

Suppose a stationary distribution $\{a_j\}$ exists. Then

$$a_j = a_j^{(n)} = \sum_i a_i \, p_{ij}^{(n)} \tag{18}$$

for each state $E_j$. Multiplying by $s^n$, with $0 \leq s < 1$, and summing,

$$a_j \sum_{n=1}^{\infty} s^n = \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} a_i \, s^n \, p_{ij}^{(n)}$$

$$= \sum_{\substack{i=1 \\ i \neq j}}^{\infty} a_i \, P_{ij}(s) + a_j(P_{jj}(s) - 1) .$$

The interchange of order of summation is valid by a standard theorem since the terms are non-negative. Thus, using (10),

$$\frac{a_j s}{1 - s} = \sum_{\substack{i=1 \\ i \neq j}}^{\infty} a_i \, F_{ij}(s) \, P_{jj}(s) + a_j \, F_{jj}(s) \, P_{jj}(s)$$

$$= P_{jj}(s) \sum_{i=1}^{\infty} a_i \, F_{ij}(s) ,$$

and hence

$$a_j s = (1 - s)\, P_{jj}(s) \sum_i F_{ij}(s)\, a_i \,. \tag{19}$$

Note that

$$F_{ij}(s) \le F_{ij}(1) = f_{ij} \le 1 \,.$$

Thus

$$\sum_i a_i\, F_{ij}(s) \le \sum_i a_i = 1 \,. \tag{20}$$

It follows that

$$\sum_i a_i\, F_{ij}(s) \tag{21}$$

converges uniformly on [0,1]. Each $F_{ij}(s)$ is continuous on [0,1], and thus $\sum_i a_i\, F_{ij}(s)$ is continuous on [0,1], and

$$\lim_{s \to 1^-} \sum_i a_i\, F_{ij}(s) = \sum_i a_i\, F_{ij}(1) \,.$$

Hence, using (17) and (19), and letting $s \to 1^-$ ,

$$a_j = \frac{1}{\mu_{jj}} \sum_i a_i\, F_{ij}(1) \,. \tag{22}$$

Using (20), it follows that

$$a_j \leq \frac{1}{\mu_{jj}} \quad .$$

For the case of a null class, $\frac{1}{\mu_{jj}} = 0$ for every state $E_j$ would imply that $a_j = 0$ for every state $E_j$. This is impossible, since

$$\sum_j a_j = 1 \quad .$$

Thus no stationary distribution exists if the irreducible Markov chain is a null class.

For the case of a positive class, clearly $f_{jj} \leq f_{ji}\, f_{ij} + (1 - f_{ji})$, $f_{jj} = 1$, and $f_{ji} > 0$. Thus $f_{ij} = 1$ and, using (22),

$$a_j = \frac{1}{\mu_{jj}} \sum_i a_i\, f_{ij} = \frac{1}{\mu_{jj}} \sum_i a_i = \frac{1}{\mu_{jj}} > 0 \quad .$$

Hence, if a stationary distribution exists it is given by

$$\{a_j\} = \{\frac{1}{\mu_{jj}}\} \quad .$$

From the relation

$$\sum_k p_{jk}^{(n)} = 1 \quad ,$$

for each $n \geq 0$, it follows that

$$\sum_{n=0}^{\infty} (s^n \sum_k p_{jk}^{(n)}) = \sum_{n=0}^{\infty} s^n = \frac{1}{1 - s} .$$

Hence

$$1 = (1 - s) \sum_{n=0}^{\infty} \sum_k s^n p_{jk}^{(n)} = (1 - s) \sum_k P_{jk}(s) ,$$

for $0 \le s < 1$. Then

$$1 = (1 - s) \sum_k (F_{jk}(s) P_{kk}(s) + p_{jk}^{(0)}) ,$$

using (10), and

$$s = (1 - s) \sum_k F_{jk}(s) P_{kk}(s) ,$$

for $0 \le s < 1$. For each $n \ge 1$ ,

$$s \ge \sum_{k=1}^{n} (1 - s) F_{jk}(s) P_{kk}(s) ,$$

and, letting $s \to 1^-$ , and using (17), it follows that

$$1 \ge \sum_{k=1}^{n} \frac{1}{\mu_{kk}} f_{jk} = \sum_{k=1}^{n} \frac{1}{\mu_{kk}} ,$$

Thus

$$1 \ge \sum_k \frac{1}{\mu_{kk}} . \tag{23}$$

By Theorem 1,

$$p_{jk}^{(n+1)} = \sum_i p_{ji}^{(n)} p_{ik} .$$

Multiplying by $s^n$, $0 < s < 1$, and summing, it follows that

$$\sum_{n=0}^{\infty} s^n p_{jk}^{(n+1)} = \sum_{n=0}^{\infty} \sum_i s^n p_{ji}^{(n)} p_{ik} ,$$

and thus

$$\frac{1}{s} P_{jk}(s) = \sum_i P_{ji}(s) p_{ik} ,$$

for $j \neq k$, where the interchange of order of summation is valid since each term is non-negative. Then, using (10),

$$\frac{1}{s} F_{jk}(s) P_{kk}(s) = \sum_i F_{ji}(s) P_{ii}(s) p_{ik} + p_{jk} .$$

Multiplying by $1 - s$, $0 < s < 1$, and letting $s \to 1^-$, it follows that

$$f_{jk} \frac{1}{\mu_{kk}} \geq \sum_{i=1}^{n} f_{ji} \frac{1}{\mu_{ii}} p_{ik} ,$$

for each $n \geq 1$. That is,

$$\frac{1}{\mu_{kk}} \geq \sum_i \frac{1}{\mu_{ii}} p_{ik} ,$$

since $f_{ji} = 1$ for every i. Noting (23),

$$\sum_k \frac{1}{\mu_{kk}} \geq \sum_k \sum_i \frac{1}{\mu_{ii}} p_{ik}$$

$$= \sum_i \frac{1}{\mu_{ii}} \left(\sum_k p_{ik}\right) = \sum_i \frac{1}{\mu_{ii}} \quad .$$

Thus,

$$\sum_k \frac{1}{\mu_{kk}} = \sum_k \sum_i \frac{1}{\mu_{ii}} p_{ik}$$

and

$$\frac{1}{\mu_{kk}} \geq \sum_i \frac{1}{\mu_{ii}} p_{ik} \quad .$$

Hence

$$\frac{1}{\mu_{kk}} = \sum_i \frac{1}{\mu_{ii}} p_{ik} \quad .$$

Now let

$$a_k' = \frac{\frac{1}{\mu_{kk}}}{\sum_j \frac{1}{\mu_{jj}}} \quad .$$

Since $1 \geq \sum_j \frac{1}{\mu_{jj}} > 0$ , $a_k' > 0$ for every k, and

$$\sum_k a_k' = 1 .$$

Also,

$$\sum_i a_i' p_{ij} = \sum_i \frac{\frac{1}{\mu_{ii}}}{\sum_j \frac{1}{\mu_{jj}}} p_{ij} = \frac{1}{\sum_j \frac{1}{\mu_{jj}}} \sum_i \frac{1}{\mu_{ii}} p_{ij}$$

$$= \frac{1}{\sum_j \frac{1}{\mu_{jj}}} \mu_{jj} = a_j' .$$

Thus $\{a_j'\}$ is a stationary distribution. Now, by earlier results, if a stationary distribution exists it is given by $\{\frac{1}{\mu_{jj}}\}$ , so $\{a_j\} = \{\frac{1}{\mu_{jj}}\}$ gives the unique stationary distribution in the case of a positive class. We have also shown that

$$\sum_j \frac{1}{\mu_{jj}} = 1 .$$

Equation (18) suggests that a proof might have been constructed by applying Theorem 18. Such is indeed the case, but the above proof, due to Levinson [1], avoids dealing with periodicity.

As a particular example, if the states of the irreducible Markov chain are ergodic, it follows by Theorem 18 that

$$\lim_{n\to\infty} p_{jk}^{(n)} = \frac{f_{jk}}{\mu_{kk}} = \frac{1}{\mu_{kk}} \quad .$$

(Notice that the effect of the initial state $E_j$ "decreases" as n increases.) Recall that

$$a_k^{(n)} = \sum_{j=1}^{\infty} a_j\, p_{jk}^{(n)} = \sum_{j=1}^{N} a_j\, p_{jk}^{(n)} + \sum_{j=N+1}^{\infty} a_j\, p_{jk}^{(n)} ,$$

for each $n \geq 0$, $N \geq 1$. Thus

$$\sum_{j=1}^{N} a_j\, p_{jk}^{(n)} \leq a_k^{(n)} \leq \sum_{j=1}^{N} a_j p_{jk}^{(n)} + \sum_{j=N+1}^{\infty} a_j ,$$

and

$$\lim_{n\to\infty} \sum_{j=1}^{N} a_j\, p_{jk}^{(n)} \leq \liminf_{n} a_k^{(n)} \leq \limsup_{n} a_k^{(n)}$$

$$\leq \lim_{n\to\infty} \sum_{j=1}^{N} a_j\, p_{jk}^{(n)} + \sum_{j=N+1}^{\infty} a_j \ .$$

That is,

$$\sum_{j=1}^{N} a_j \frac{1}{\mu_{kk}} \leq \liminf_{n} a_k^{(n)} \leq \limsup_{n} a_k^{(n)}$$

$$\leq \sum_{j=1}^{N} a_j \frac{1}{\mu_{kk}} + \sum_{j=N+1}^{\infty} a_j \quad .$$

But,

$$\sum_{j=N+1}^{\infty} a_j \to 0 \quad \text{as} \quad N \to \infty$$

and thus

$$\frac{1}{\mu_{kk}} \le \liminf_{n} a_k^{(n)} \le \limsup_{n} a_k^{(n)} \le \frac{1}{\mu_{kk}} .$$

Hence

$$\lim_{n\to\infty} a_k^{(n)} = \frac{1}{\mu_{kk}} .$$

Recalling the remark immediately preceding Theorem 21, we see that for this example the stationary distribution $\{\frac{1}{\mu_{kk}}\}$ is approached in the limit, regardless of the initial distribution $\{a_k\}$ . It should be emphasized that at this point we are considering only irreducible Markov chains.

If the states of the irreducible Markov chain are positive and have period t, then states belonging to the same cyclic subclass form an irreducible Markov chain with transition matrix $P' = P^t$, as described preceding Theorem 20, and with period one. Thus, as in the example above,

$$\lim_{n\to\infty} p'^{(n)}_{jk} = \frac{1}{\mu'_{kk}} ,$$

or, equivalently,

$$\lim_{n\to\infty} p_{jk}^{(nt)} \to \frac{t}{\mu_{kk}} \quad .$$

(Recall the last part of the proof of Theorem 17. The above result is actually Theorem 18, with $f_{jk}(r) = 0$ for $r \neq t$, $f_{jk}(t) = 1$.) Here,

$$\frac{1}{\mu'_{kk}} = \frac{t}{\mu_{kk}} \quad ,$$

yields the unique stationary distribution for the subchain, and is approached in nt steps as $n \to \infty$ , regardless of the initial distribution of the subchain. Note that the stationary distribution for the entire chain is $\{\frac{1}{\mu_{kk}}\}$ , by Theorem 21.

Theorem 22. For any positive class C,

$$\sum_{E_j \varepsilon C} \frac{1}{\mu_{jj}} = 1 \quad .$$

Proof. Any positive class is closed, and so may be considered as an irreducible Markov chain. Since $\{\frac{1}{\mu_{jj}}\}$ is a probability distribution for that chain,

$$\sum_{E_j \varepsilon C} \frac{1}{\mu_{jj}} = 1 \ .$$

Note that the value of $\mu_{jj}$ does not depend on whether we consider the original chain or the subchain.

The above results on cyclic subclasses are to be expected intuitively, for if we consider the entire chain on some arbitrary step n, for large n, the process is equally likely to be in any one of the t cyclic subclasses. The probability of being in C(r), r = 0, 1, ..., t-1, is thus $\frac{1}{t}$. We have seen that the probability of being in state $E_j$ is $\frac{1}{\mu_{jj}}$, considering the entire chain, or $\frac{t}{\mu_{jj}}$ if we have the additional information that we are in the cyclic subclass of $E_j$. Thus follows the obvious relation,

$$\frac{1}{T}\frac{t}{\mu_{jj}} = \frac{1}{\mu_{jj}} \quad ,$$

that is,

$$P\{x_n(\omega)\varepsilon C_j\}P\{x_n = E_j | x_n(\omega)\varepsilon C_j\} = P\{x_n = E_j\} \quad ,$$

where $C_j$ is the cyclic subclass of state $E_j$. Further, the assertions

$$\sum_k \frac{1}{\mu_{kk}} = 1 \quad ,$$

and

$$\sum_{E_k \varepsilon C(r)} \frac{t}{\mu_{kk}} = 1 \, ,$$

are in agreement since

$$\sum_k \frac{1}{\mu_{kk}} = \frac{1}{t} \sum_{r=0}^{t-1} \sum_{E_k \varepsilon C(r)} \frac{t}{\mu_{kk}} = \frac{1}{t} t = 1 \quad .$$

Theorem 21 may be extended to apply to Markov chains which are not necessarily irreducible, using an idea similar to that of the above example. (Recall the division of the not necessarily irreducible Markov chain into classes preceding Theorem 20.)

Theorem 23. In any Markov chain, if all the states are null then there is no stationary distribution. If there exist positive states then the only stationary distribution is that defined by $a_j = 0$ for all null states $E_j$, and

$$a_j = \frac{p_t}{\mu_{jj}}$$

for all states $E_j$ belonging to positive class $C_t$, where the $p_t$ are arbitrary non-negative numbers such that

$$\sum_t p_t = 1 \ .$$

Proof. From the proof of Theorem 21, if

$$a_j = \sum_i a_i \, p_{ij}^{(n)}$$

then $a_j = 0$ for all null states $E_j$. Thus, if all states are null,

$$\sum_j a_j = 0 \ ,$$

in which case $\{a_j\}$ is not a probability distribution, and there is no stationary distribution.

If there exist positive states they are in closed classes $C_t$. Considering each of these classes as an irreducible Markov chain and using Theorem 21,

$$\{a_j'\} = \{\frac{1}{\mu_{jj}}\}$$

is the unique stationary distribution in the subchain. If $p_t$ is the probability of being in class $C_t$, then

$$\{a_j\} = \{\frac{p_t}{\mu_{jj}}\}$$

gives a stationary distribution for the entire chain, since

$$\sum_j a_j = \sum_t \sum_{E_j \varepsilon C_t} \frac{p_t}{\mu_{jj}} = \sum_t p_t \sum_{E_j \varepsilon C_t} \frac{1}{\mu_{jj}}$$

$$= \sum_t p_t = 1 \ ,$$

using Theorem 22, and for $E_j \varepsilon C_t$ ,

$$\sum_i a_i \, p_{ij} = \sum_{E_i \varepsilon C_t} a_i \, p_{ij} = \sum_{E_i \varepsilon C_t} \frac{p_t}{\mu_{ii}} p_{ij}$$

$$= p_t \sum_{E_i \varepsilon C_t} a_i' \, p_{ij} = p_t \, a_j' = \frac{p_t}{\mu_{jj}} = a_j \ ,$$

while

$$\sum_i a_i \, p_{ij} = \sum_{E_i \text{ positive}} a_i \, p_{ij} = 0 = a_j \ ,$$

for $E_j$ null. On the other hand, if $\{a_j\}$ is a stationary distribution then, from the above results, $a_j = 0$ for all null states $E_j$. If $E_j \varepsilon C_t$, a positive class, and $E_i \notin C_t$, then either $E_i$ is null, in which case $a_i = 0$, or $E_i \varepsilon C_k$, $k \neq t$, a positive class, in which case $p_{ij}^{(n)} = 0$ for all $n \geq 0$. Thus,

$$a_j = \sum_i a_i \, p_{ij}^{(n)} = \sum_{E_i \varepsilon C_t} a_i \, p_{ij}^{(n)}$$

$$= \frac{1}{n} \Big( \sum_{E_i \varepsilon C_t} a_i \, p_{ij}^{(1)} + \ldots + \sum_{E_i \varepsilon C_t} a_i \, p_{ij}^{(n)} \Big)$$

$$= \sum_{E_i \varepsilon C_t} a_i \Big( \frac{1}{n} \sum_{m=1}^{n} p_{ij}^{(m)} \Big) \to \sum_{E_i \varepsilon C_t} a_i \, \frac{f_{ij}}{\mu_{jj}} \ ,$$

as $n \to \infty$, by Theorem 18, where the limit is taken inside the summation by a proof exactly like that in the example following Theorem 21. Also, from the proof of Theorem 21, $f_{ij} = 1$, and thus

$$a_j = \frac{1}{\mu_{jj}} \sum_{E_i \varepsilon C_t} a_i \ .$$

But

$$\sum_{E_i \varepsilon C_t} a_i = p_t \; ,$$

and

$$\sum_t \sum_{E_i \varepsilon C_t} a_i = \sum_i a_i = 1 \; .$$

Theorem 23 may be used to classify states as null or positive by determining whether or not a stationary distribution exists and whether or not $a_j$ must be zero.

## Transient States

We now focus attention on transient states. If $E_j$ is a transient state and state $E_k$ is positive, with period $t_k$, then by Theorem 18,

$$\lim_{n\to\infty} p_{jk}^{(nt_k+r)} = f_{jk}(r) \frac{t_k}{\mu_{kk}} \; .$$

If $E_i$ is any other state in the same cyclic subclass as $E_k$, then $f_{ki} = 1$ from the proof of Theorem 21, and

$$f_{jk}(r) \geq f_{ji}(r)\, f_{ik} = f_{ji}(r) \; .$$

Similarly, $f_{ji}(r) \geq f_{jk}(r)$, and thus

$$f_{jk}(r) = f_{ji}(r) .$$

That is, the probability $f_{jk}(r)$ is the same for all states in the cyclic subclass of state $E_k$. (Also, if state $E_{k(1)}$ belongs to the next cyclic subclass,

$$1 = f_{kk(1)} = \sum_{r=1}^{t_k} f_{kk(1)}(r) = f_{kk(1)}(1) ,$$

and thus

$$f_{jk}(r) = f_{jk}(r) f_{kk(1)}(1) \leq f_{jk(1)}(r + 1) ,$$

and the inequalities continue around the cycle,

$$f_{jk}(r) \leq f_{jk(1)}(r + 1) \leq \ldots \leq f_{jk(t_k)}(r + t_k) = f_{jk}(r) ,$$

showing that the inequalities become equalities.) Thus, if $E_j$ is transient and $C(m)$ is a cyclic subclass of positive class $C$, $m = 1, \ldots, t_k$, then for each state $E_k$ of $C(m)$, as $n \to \infty$ ,

$$p_{jk}^{(nt_k+r)} \to x_{jm}(r) \frac{t_k}{\mu_{kk}} ,$$

where $x_{jm}(r)$ is the probability of passage from $E_j$ to $C(m)$ in $r$ (mod $t_k$) steps. Then the probability that starting from transient state $E_j$ the system will ever enter $C(m)$, and hence also $C$, the entire positive class, is

$$\sum_{r=1}^{t_k} x_{jm}(r) = x_j \quad .$$

Note also that

$$\sum_{m=1}^{t_k} x_{jm}(r) = x_j \quad .$$

Define $x_j^{(n)}$ to be the probability that at step n, and not before, the system reaches a state in C (and then stays in C, since C is closed), having begun in transient state $E_j$. Then,

$$\sum_{n=1}^{\infty} x_j^{(n)} = x_j \ .$$

The number $x_j$ is the probability of absorption in C from transient state $E_j$. We will let T denote the set of all transient states in the chain and suppose that the system is initially in the transient state $E_j$. The probability of absorption in other closed sets, or staying forever in T, is $1 - x_j$. (It is possible to stay forever in T, as, for example, if the states are $E_1$, $E_2$, ..., with $p_{i,i+1} = 1$, for $i \geq 1$.)

We have

$$x_j^{(1)} = \sum_{E_k \varepsilon C} p_{jk} \ ,$$

(and in general

$$\sum_{m=1}^{n} x_j^{(m)} = \sum_{E_k \varepsilon C} p_{jk}^{(n)} \quad ,$$

for $n \geq 1$), and

$$x_j^{(n+1)} = \sum_{E_i \varepsilon T} p_{ji}\, x_i^{(n)} \quad ,$$

for $n \geq 1$. These two relationships completely determine the numbers $x_j^{(n)}$, for $E_j \varepsilon T$.

Then,

$$\sum_{n=1}^{\infty} x_j^{(n+1)} = \sum_{n=1}^{\infty} \sum_{E_i \varepsilon T} p_{ji}\, x_i^{(n)}$$

$$= \sum_{E_i \varepsilon T} p_{ji} \left( \sum_{n=1}^{\infty} x_i^{(n)} \right)$$

$$= \sum_{E_i \varepsilon T} p_{ji}\, x_i \; .$$

That is,

$$x_j - x_j^{(1)} = \sum_{E_i \varepsilon T} p_{ji}\, x_i \quad ,$$

or

$$x_j^{(1)} = x_j - \sum_{E_i \varepsilon T} p_{ji}\, x_i \ .$$

The $x_j$, for $E_j \varepsilon T$, satisfy the system of linear equations

$$z_j - \sum_{E_i \varepsilon T} p_{ji}\, z_i = \sum_{E_k \varepsilon C} p_{jk} \ . \tag{24}$$

Thus, given a transient state $E_j$ and a persistent closed set C, positive or null, the probability $x_j$ that starting from $E_j$ the system will ever enter C is given as a member of the solution of the system of equations (24). We shall consider uniqueness of a solution of (24) shortly.

Let $y_j^{(n)}$ be the probability that the system is in a transient state at step n, having started at transient state $E_j$. Then

$$y_j^{(1)} = \sum_{E_i \varepsilon T} p_{ji} \ ,$$

and

$$y_j^{(n+1)} = \sum_{E_i \varepsilon T} p_{ji}\, y_i^{(n)} \ , \tag{25}$$

since once having left the transient states there can be no return. We know

$$1 = \sum_k p_{jk} \geq \sum_{E_k \varepsilon T} p_{jk} = y_j^{(1)} \ ,$$

$$y_j^{(2)} = \sum_{E_k \varepsilon T} p_{jk}\, y_k^{(1)} \le \sum_{E_k \varepsilon T} p_{jk} = y_j^{(1)} \quad ,$$

and, in general, if $y_j^{(n)} \le y_j^{(n-1)}$, for every $E_j \varepsilon T$, then

$$y_j^{(n+1)} = \sum_{E_i \varepsilon T} p_{ji}\, y_i^{(n)} \le \sum_{E_i \varepsilon T} p_{ji}\, y_i^{(n-1)} = y_j^{(n)} \quad .$$

Thus $\{y_j^{(n)}\}$ is a non-increasing sequence of non-negative numbers for each state $E_j \varepsilon T$. Hence

$$y_j = \lim_{n \to \infty} y_j^{(n)}$$

exists. The number $y_j$ is the probability of forever staying in T, having started in transient state $E_j$.

Denoting the transient states by $E_1$, $E_2$, ... (if there are only a finite number of transient states the following results will be obvious), we have

$$y_j^{(n+1)} = \sum_{k=1}^{\infty} p_{jk}\, y_k^{(n)} = \sum_{k=1}^{m} p_{jk}\, y_k^{(n)} + \sum_{k=m+1}^{\infty} p_{jk}\, y_k^{(n)} \quad ,$$

and hence

$$\sum_{k=1}^{m} p_{jk}\, y_k^{(n)} \le y_j^{(n+1)} \le \sum_{k=1}^{m} p_{jk}\, y_k^{(n)} + \sum_{k=m+1}^{\infty} p_{jk} \; ,$$

since $0 \leq p_{jk}, y_k^{(n)} \leq 1$. First letting $n \to \infty$, and then $m \to \infty$, we have

$$\sum_{E_k \varepsilon T} p_{jk}\, y_k = y_j \quad .$$

The $y_j$, for $E_j \varepsilon T$, satisfy the system of linear equations

$$z_j = \sum_{E_k \varepsilon T} p_{jk}\, z_k \ . \tag{26}$$

Suppose there is another bounded solution (i.e, a solution $\{u_j\}$ such that $|u_n| \leq C$ for some real C, and all n) $\{u_j\}$, and assume $|u_j| \leq 1$ since $\{cu_j\}$ is also a solution for any constant c. Then

$$y_j^{(1)} = \sum_{E_k \varepsilon T} p_{jk} \geq \sum_{E_k \varepsilon T} p_{jk}|u_k| \geq |u_j| \ .$$

It follows easily by induction, using (25), that

$$y_j^{(n)} \geq |u_j| \ ,$$

for all $n \geq 1$, and hence

$$y_j \geq |u_j| \quad .$$

Thus, $\{y_j\}$ is the maximal solution of (26), bounded by one, so if $y_j = 0$ for all $E_j \varepsilon T$, then the only bounded solution is the zero solution (recall that if $\{u_j\}$ is a solution such that $|u_j| \leq m$, where $m > 0$, then $\{\frac{u_j}{m}\}$ is a solution with $|\frac{u_j}{m}| \leq 1$). Conversely, since $\{y_j\}$ is a bounded solution, if (26) has no non-zero bounded solutions then $y_j = 0$ for all $E_j \varepsilon T$. Thus, $y_j = 0$ for all $E_j \varepsilon T$ if and only if (26) has no non-zero bounded solutions.

Consider again the system (24). We know that the $x_j$ yield a bounded solution. Suppose $\{w_j\}$ is also a bounded solution. Then, subtracting,

$$x_j - w_j = \sum_{E_i \varepsilon T} p_{ji}\, x_i - \sum_{E_i \varepsilon T} p_{ji} w_i = \sum_{E_i \varepsilon T} p_{ji}\, (x_i - w_i) .$$

Thus $\{x_j - w_j\}$ is a bounded solution of (26). Thus, $x_j = w_j$ for all $E_j \varepsilon T$ if $y_j = 0$ for all $E_j \varepsilon T$, and $\{x_j\}$ is the unique bounded solution of (24) in this case. In case $y_j > 0$ for some $E_j \varepsilon T$, then (26) has a bounded non-zero solution and

$$\begin{aligned}(x_j - y_j) - \sum_{E_i \varepsilon T} p_{ji}(x_i - y_i) &= x_j - \sum_{E_i \varepsilon T} p_{ji}\, x_i - y_j + \sum_{E_i \varepsilon T} p_{ji}\, y_i \\ &= x_j - \sum_{E_i \varepsilon T} p_{ji}\, x_i \\ &= \sum_{E_k \varepsilon C} p_{jk} \quad .\end{aligned}$$

Hence $\{x_j - y_j\}$ is a bounded solution of (24) distinct from $\{x_j\}$, and in this case the solution is not unique. We summarize the above results in the following theorem.

Theorem 24. The sequence of probabilities $\{y_j\}$ is a maximal solution, bounded by one, of (26), and $y_j = 0$ for all $E_j \varepsilon T$ if and only if (26) has no non-zero bounded solutions.

The sequence of probabilities $\{x_j\}$ is a solution of (24). This is the unique bounded solution of (24) except when there exists a $y_j > 0$.

In general we cannot say that $y_j$ must equal zero. For example, in the Markov chain example earlier in this section, $y_j = 1$ for every state $E_j$.

Theorem 25. In a finite Markov chain the probability of staying forever in the transient states is zero. The $x_j$ are determined as the unique solution of (24).

Proof. From the discussion following Theorem 20, the states may only be inessential or positive, and at least one state must be positive. Let M be the maximum of the finite number of probabilities $y_j$, ordered so that $M = y_1 = y_2 = \ldots = y_a$, and $M > y_{a+1} \geq \ldots \geq y_n$, $1 \leq a \leq n$, where the transient states are $E_1, E_2, \ldots E_n$, and suppose that $M > 0$. Then for $j \leq a$,

$$M = y_j = \sum_{E_i\, T} p_{ji} y_i = \sum_{i=1}^{a} p_{ji}\, M + \sum_{i=a+1}^{n} p_{jk}\, y_i \quad .$$

Thus

$$\sum_{i=1}^{a} p_{ji} = 1 \ .$$

Hence, $E_1, E_2, \ldots, E_a$ form a closed set of states, and this is impossible since they are inessential. Thus, $M = 0$, in which case $y_j = 0$ for every state $E_j \varepsilon T$, and the solution of (24) is unique, by Theorem 24.

## Theorems on the Classification of States

Following Theorem 23 it was mentioned that that theorem could be used to classify states as null or positive. The following theorems are also useful in classifying states as transient, persistent null, or positive.

Theorem 26. Suppose an irreducible Markov chain has states $E_o, E_1, E_2, \ldots$ . In order that the states be transient it is necessary and sufficient that the system of equations

$$y_i = \sum_{j=s}^{\infty} p_{ij} y_j ,$$

for $i = s, s+1, \ldots, s \geq 1$, admits of a non-zero bounded solution.

Proof. Let $T$ be the set of states $E_s, E_{s+1}, \ldots$, and $y_j^{(n)}$ be the probability that the system is in $T$ at step $n$, never having been out of $T$, having started in state $E_j \varepsilon T$. Then, exactly as before, $y_j$ is the probability of never leaving $T$, and $y_i = 0$ for all states $E_i \varepsilon T$, if and only if

$$y_i = \sum_{j=s}^{\infty} p_{ij} y_j ,$$

$i = s, s+1, \ldots$, has no non-zero bounded solution. If $y_i = 0$ for all states $E_i \varepsilon T$, then the probability of entering the set of states $E_o, E_1, \ldots, E_{s-1}$, is one, from which it follows that

$$(1 - f_{io})(1 - f_{i1}) \ldots (1 - f_{i,s-1}) = 0 ,$$

for i = 0, 1, ..., s - 1. Thus, using

$$f_{ij}\, f_{jk} \cdots f_{mp} \le f_{ip} ,$$

we have $f_{ii} = 1$, for some i = 0, 1, ..., s - 1, in which case the state $E_i$ is persistent. Hence all states are persistent. If there is a non-zero bounded solution, then $y_i > 0$ for some state $E_i \varepsilon T$, since $\{y_i\}$ is the maximal solution bounded by one. Since the Markov chain is irreducible, $f_{oi} > 0$. But $f_{io} < 1$, and thus

$$f_{oo} \le f_{oi}\, f_{io} + (1 - f_{oi}) < 1 .$$

Hence $E_o$ is transient, and consequently all states are transient.

The remainder of this section follows Foster [1].

Theorem 27. An irreducible Markov chain with states $E_o$, $E_1$, ..., is positive if there exists a non-zero solution of the equations

$$\sum_i v_i p_{ij} = v_j , \tag{27}$$

for j = 0, 1, ..., such that

$$\sum_i |v_i| < \infty \; ; \tag{28}$$

and only if (28) holds for any solution of the inequalities

$$\sum_i v_i \, p_{ij} \leq v_j \; , \tag{29}$$

$j = 0, 1, \ldots$, for which $v_i \geq 0$, for $i \geq s \geq 0$, and a fixed s.

<u>Proof</u>. Let $\{v_j\}$ be a non-zero solution of (27), satisfying (28). Then

$$\sum_i v_i \, p_{ij}^{(2)} = \sum_i \sum_k v_i \, p_{ik} \, p_{kj} = \sum_k \sum_i v_i \, p_{ik} \, p_{kj}$$

$$= \sum_k v_k \, p_{kj} = v_j \; ,$$

where the interchange of the order of summation is valid by a standard theorem (for example, Apostol [1], Theorem 13-9). It follows easily by induction that

$$v_j = \sum_i v_i \, p_{ij}^{(n)} \quad ,$$

for $j = 0, 1, \ldots$, and all $n \geq 1$. Now note that

$$\sum_{i=0}^{m} v_i \, p_{ij}^{(n)} - \sum_{i=m+1}^{\infty} |v_i| \leq \sum_{i=0}^{\infty} v_i \, p_{ij}^{(n)} \leq \sum_{i=0}^{m} v_i \, p_{ij}^{(n)} + \sum_{i=m+1}^{\infty} |v_i| \; .$$

Hence

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{i=0}^{m} v_i\, p_{ij}^{(n)} - \sum_{i=m+1}^{\infty}|v_i| \le v_j \le \frac{1}{N}\sum_{n=1}^{N}\sum_{i=0}^{m} v_i\, p_{ij}^{(n)} + \sum_{i=m+1}^{\infty}|v_i| \; .$$

Now let $N \to \infty$, and then $m \to \infty$, and apply Theorem 18. It follows that

$$\sum_{i=0}^{\infty} v_i \frac{f_{ij}}{\mu_{jj}} = v_j \qquad (j = 0, 1, \ldots) \; ,$$

and $\frac{1}{\mu_{jj}} > 0$ for all $j$ or $0$ for all $j$.

Since the chain is irreducible, each $f_{ij} = 1$. Thus

$$\frac{1}{\mu_{jj}} \sum_{i=0}^{\infty} v_i = v_j \qquad (j = 0, 1, \ldots) \; . \tag{30}$$

Since $\{v_j\}$ is a non-zero solution, it follows from (30) that at least one component $v_j > 0$. Hence $\frac{1}{\mu_{jj}} > 0$ for all $j$. Hence the chain is positive.

Conversely, if the states are positive, suppose that $\{v_i\}$ is a solution of (29) such that $v_i \ge 0$ for all $i \ge s \ge 0$ (for some $s \ge 0$). A simple calculation shows that

$$\sum_j \sum_i v_i\, p_{ij}\, p_{jk} \le \sum_j v_j\, p_{jk} \; .$$

Thus, noting (29) once again, and reversing the order of summation, it follows that

$$\sum_i v_i \, p_{ik}^{(2)} \leq \sum_j v_j \, p_{jk} \leq v_k \qquad (k = 0, 1, \ldots) .$$

An easy induction proof shows that

$$\sum_i v_i \, p_{ik}^{(n)} \leq v_k \qquad (k = 0, 1, \ldots) \tag{31}$$

for each $n \geq 1$. For each $m \geq s$ and $N \geq 1$,

$$\frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{m} v_i \, p_{ik}^{(n)} \leq v_k . \tag{32}$$

Since the chain is irreducible and positive, an application of Theorem 18 yields

$$\sum_{i=1}^{m} v_i \frac{1}{\mu_{kk}} \leq v_k .$$

Hence

$$\sum_{i=1}^{\infty} v_i \leq \mu_{kk} \, v_k < \infty .$$

Since $v_i \geq 0$ for all $i \geq s$, it follows that the series is absolutely convergent, and thus

$$\sum_{i=1}^{\infty} |v_i| < \infty .$$

(Note from (32) if the states are null instead of positive, each $v_k \geq 0$.)

<u>Theorem 28</u>. An irreducible Markov chain with states $E_o$, $E_1$, ..., is positive if there exists a non-negative solution of the inequalities

$$\sum_{j=0}^{\infty} p_{ij}\, y_j \leq y_i - 1 ,$$

$i \neq 0$, such that

$$\sum_{j=0}^{\infty} p_{oj}\, y_j < \infty .$$

<u>Proof</u>. Let $\{y_i\}$ be such a solution, and define

$$y_i^{(1)} = y_i, \quad y_i^{(n+1)} = \sum_{j=0}^{\infty} p_{ij}^{(n)}\, y_j ,$$

and

$$\lambda = \sum_{j=0}^{\infty} p_{oj}\, y_j \geq 0 .$$

Then

$$y_i^{(n+1)} = \sum_{j=0}^{\infty} p_{ij}^{(n)}\, y_j = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} p_{ik}^{(n-1)}\, p_{kj}\, y_j$$

$$= \sum_{k=0}^{\infty} p_{ik}^{(n-1)} y_k^{(2)} = p_{io}^{(n-1)} y_o^{(2)} + \sum_{k=1}^{\infty} p_{ik}^{(n-1)} y_k^{(2)}$$

$$\leq p_{io}^{(n-1)} \lambda + \sum_{k=1}^{\infty} p_{ik}^{(n-1)} (y_k - 1)$$

$$= p_{io}^{(n-1)} \lambda + y_1^{(n)} - p_{io}^{(n-1)} y_o - (1 - p_{io}^{(n-1)})$$

$$\leq (\lambda + 1) p_{io}^{(n-1)} + y_i^{(n)} - 1 .$$

Since $\lambda$ is finite and $y_i^{(1)}$ is finite for every i, $y_i^{(n)}$ is finite for every i and all $n \geq 1$. Also, for $n \geq 1$,

$$0 \leq y_o^{(n+2)} \leq (\lambda + 1) \sum_{m=1}^{n} p_{oo}^{(m)} + y_o^{(2)} - n ,$$

and thus

$$0 \leq (\lambda + 1) \frac{1}{n} \sum_{m=1}^{n} p_{oo}^{(m)} + \frac{y_o^{(2)}}{n} - 1 .$$

Letting $n \to \infty$ , and using Theorem 18, it follows that

$$0 \leq (\lambda + 1) \frac{f_{oo}}{\mu_{oo}} - 1 ,$$

or,

$$\frac{f_{oo}}{\mu_{oo}} \geq \frac{1}{\lambda + 1} > 0 .$$

It follows that $\mu_{oo}$ is finite. Thus state $E_o$ is positive, and hence the chain is positive.

Theorem 29. If an irreducible Markov chain with states $E_o$, $E_1$, ..., is positive, then the expected passage times $\mu_{jo}$ satisfy the equations

$$\sum_{j=1}^{\infty} p_{ij}\, \mu_{jo} = \mu_{io} - 1 ,$$

for $i \geq 0$. (This theorem could, of course, be stated for $\mu_{jk}$, for any state $E_k$, since the labeling is arbitrary.)

Proof. State $E_i$ is positive, and thus $\mu_{ii}$ is finite for all i. For $i \neq j$, let

$$R_{ij}^{(n)} = P\{x_n = E_j,\ x_k \neq E_i \text{ for } 1 \leq k \leq n \mid x_o = E_i\} ,$$

for $n \geq 1$. Then

$$f_{ii}^{(v)} = \sum_{j \neq i} R_{ij}^{(n)}\, f_{ji}^{(v-n)} ,$$

for $1 \leq n < v$. It follows that

$$\infty > \mu_{ii} = \sum_{v=1}^{\infty} v\, f_{ii}^{(v)} = \sum_{v=1}^{n} v\, f_{ii}^{(v)} + \sum_{v=n+1}^{\infty} v \sum_{j \neq i}^{\infty} R_{ij}^{(n)}\, f_{ji}^{(v-n)}$$

$$= \sum_{v=1}^{n} v\, f_{ii}^{(v)} + \sum_{j \neq 1} \sum_{v=1}^{\infty} R_{ij}^{(n)}\, (v + n)\, f_{ji}^{(v)}$$

$$= \sum_{v=1}^{n} v\, f_{ii}^{(v)} + \sum_{j \neq i} R_{ij}^{(n)} (\mu_{ji} + n\, f_{ji})$$

(since

$$\sum_{v=1}^{\infty} R_{ij}^{(n)} (v + n)\, f_{ji}^{(v)} \ , \quad \sum_{v=1}^{\infty} n\, R_{ij}^{(n)}\, f_{ji}^{(v)} \ ,$$

are both finite, and, hence, their difference,

$$\sum_{v=1}^{\infty} R_{ij}^{(n)}\, v\, f_{ji}^{(v)} \ ,$$

is finite)

$$= \sum_{v=1}^{n} v\, f_{ii}^{(v)} + \sum_{j \neq i} R_{ij}^{(n)} (\mu_{ji} + n) \ .$$

Also,

$$p_{ij}^{(v)} = \sum_{n=0}^{v-1} p_{ii}^{(n)}\, R_{ij}^{(v-n)} \ ,$$

and $p_{ij}^{(v)} > 0$ for some $v \geq 1$, and thus $R_{ij}^{(n)} > 0$ for some $n \geq 1$. Thus

$$\mu_{ji} < \infty \ ,$$

for i, j = 0, 1, ... .

Then

$$\infty > \mu_{io} - 1 = \mu_{io} - \sum_{n=1}^{\infty} f_{io}^{(n)}$$

$$= \mu_{io} - f_{io}^{(1)} - \sum_{n=1}^{\infty} f_{io}^{(n+1)}$$

$$= \sum_{n=1}^{\infty} (n + 1)\, f_{io}^{(n+1)} - \sum_{n=1}^{\infty} f_{io}^{(n+1)}$$

$$= \sum_{n=1}^{\infty} n\, f_{io}^{(n+1)}$$

$$= \sum_{n=1}^{\infty} n \,\left(\sum_{j=1}^{\infty} p_{ij}\, f_{jo}^{(n)}\right)$$

$$= \sum_{j=1}^{\infty} p_{ij}\, \mu_{jo} \,,$$

for $i \geq 0$.

<u>Theorem 30</u>. An irreducible Markov chain with states $E_o$, $E_1$, ..., is persistent if there exists a solution $\{y_i\}$ of the inequalities

$$\sum_{j=0}^{\infty} p_{ij}\, y_j \leq y_i \,, \tag{33}$$

for $i \neq 0$, such that $y_i \to \infty$ as $i \to \infty$ .

<u>Proof</u>. (Following Kendall [1] as well as Foster [1].) Consider the modified Markov chain with $p'_{oo} = 1$, and $p'_{ij} = p_{ij}$ for $i \neq 0$. Then

state $E_o$ is positive and all other states are inessential (since $p'_{io}{}^{(n)} \geq p_{io}{}^{(n)} > 0$ for some $n \geq 1$, $i \neq 0$, but $p'_{oi}{}^{(n)} = 0$ for every $n \geq 1$), thus transient by Theorem 16, thus null. By Theorem 18,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p'_{ij}{}^{(m)} = 0 \ , \tag{34}$$

for $j \neq 0$, and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p'_{io}{}^{(m)} = \frac{f'_{io}}{\mu'_{oo}} = f'_{io} \ . \tag{35}$$

Let $\{y_i\}$ be such a solution of (33). Since $\{y_i+c\}$ for an arbitrary constant c is also such a solution, and only finitely many $y_i < 1$ since $y_i \to \infty$ , we may assume that $y_i \geq 1$ for all i. Then

$$\sum_{j=0}^{\infty} p'_{ij}\, y_j = \sum_{j=0}^{\infty} p_{ij}\, y_j \leq y_i \ ,$$

for $i \neq 0$, and if

$$\sum_{j=0}^{\infty} p'_{ij}{}^{(n)}\, y_j \leq y_i \ , \tag{36}$$

for $i \neq 0$, then, since (36) is trivially valid for $i = 0$,

$$\sum_{j=0}^{\infty} p'_{ij}{}^{(n+1)}\, y_j = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} p'_{ik}\, p'_{kj}{}^{(n)}\, y_j$$

$$= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} p'_{ik} \, p'^{(n)}_{kj} \, y_j \le \sum_{k=0}^{\infty} p'_{ik} \, y_k \le y_i \quad .$$

Thus, by induction, (36) is valid for $n \geq 1$. Let

$$v_m = \mathrm{Min}\{y_{m+1}, y_{m+2}, \ldots\} \geq 1 \ .$$

Then, for $n \geq 1$,

$$\sum_{j=m+1}^{\infty} p'^{(n)}_{ij} \, v_m \le y_i \ ,$$

$$v_m(1 - \sum_{j=0}^{m} p'^{(n)}_{ij}) \le y_i \quad ,$$

and

$$1 - \frac{y_i}{v_m} \le \sum_{j=0}^{m} p'^{(n)}_{ij} \quad .$$

Thus,

$$1 - \frac{y_i}{v_m} \le \sum_{j=0}^{m} \sum_{k=1}^{n} \frac{p'^{(k)}_{ij}}{n} \quad .$$

Now, letting $n \to \infty$ and then $m \to \infty$ , using (34), (35), and the fact that $v_m \to \infty$ as $m \to \infty$ , it follows that $1 \le f'_{io}$ . Hence, $f'_{io} = 1$ for every i. But for $i \neq 0$, $f'_{io} = f_{io}$, and thus $f_{io} = 1$ for $i \neq 0$. Then using

$$1 - f_{oo} = \sum_{j=1}^{\infty} p_{oj}(1 - f_{jo}) \ , \tag{37}$$

the probability of not returning to state $E_o$, it follows that $f_{oo} = 1$. Thus the state $E_o$ is persistent. Hence the original Markov chain is persistent.

Theorem 31. An irreducible Markov chain with states $E_o$, $E_1$, ..., is transient if and only if there exists a bounded solution $\{y_i\}$ of

$$\sum_{j=0}^{\infty} p_{ij}\, y_j \le y_i \tag{38}$$

for $i \neq 0$, such that $y_i < y_o$ for some i.

Proof. Using the modified chain as above,

$$\sum_{j=0}^{\infty} p'_{ij}\, f'_{jo} = f'_{io} \ ,$$

and for $i \neq 0$, $p'_{ij} = p_{ij}$. Hence

$$\sum_{j=0}^{\infty} p_{ij}\, f'_{jo} = f'_{io} \ .$$

Thus, $\{f'_{io}\}$ is a bounded solution of (38), with $y_o = f'_{oo} = 1$. If the original chain is transient, $y_i = f'_{io} < 1$ for some i, using (37), and thus $y_i < y_o$ for some i. Conversely, let $\{y_i\}$ be such a bounded solution of (38). Then for $\beta > 0$, and $\alpha$ arbitrary,

$$\sum_{j=0}^{\infty} p_{ij}(\alpha + \beta y_j) = \alpha + \beta \sum_{j=0}^{\infty} p_{ij}\, y_j \le \alpha + \beta y_i$$

for $i \neq 0$, $\{\alpha + \beta y_i\}$ is bounded, and $\alpha + \beta y_i < \alpha + \beta y_o$ for some i, and hence we may assume that $y_o = 1$ and $0 \le y_i \le 2$. Then (36) holds, and hence

$$p_{io}^{'(n)}\, y_o \le y_i$$

for every i and n. Thus

$$\frac{1}{n} \sum_{m=1}^{n} p_{io}^{'(m)} \le y_i \ .$$

Letting $n \to \infty$ , and using Theorem 18, it follows that

$$f_{io}^{'} \le y_i$$

for each i. But $y_i < y_o = 1$ for some i, and thus, $f_{io}^{'} < 1$ for some i. Thus, using (37), $f_{oo} < 1$, and the Markov chain is transient.

# CHAPTER III

## SOME ASPECTS OF QUEUEING THEORY

This chapter contains an introduction to the basic concepts of queueing theory, but is presented here primarily to illustrate in detail how the results of Chapter II may be applied to certain special queueing systems.

The queueing problem is this: At a certain location, usually called a counter, customers arrive seeking service from a specified number of servers. If at some time there are more customers than servers, some of the customers will have to form a queue and wait until a server becomes available. Of particular interest is the probability of a specified number of customers in the queue at a specified time and the asymptotic behavior of that probability. Queueing systems will be classified according to the arrival pattern, service mechanism, and queue discipline.

We shall present the method of the imbedded Markov chain, and consider two special queueing systems to which the method is applicable. In the first system the arrivals are random, the distribution of the service times is not specified (but the service time distributions are assumed to be identical and independent), and there is but one server. In the second system the arrival distribution is not specified (but the inter-arrival times are assumed to be identically and independently distributed), the service times are random, and there are a finite number of servers. The stochastic process associated with each of these systems is, in general, not Markovian, but we are able to define an imbedded Markov chain so that the results of Chapter II may be used. In each system the imbedded Markov chain is aperiodic and irreducible, and is positive, persistent null, or transient according as the relative traffic intensity $\rho$ (the ratio of average input to average output of the system) is less than, equal to, or greater than one.

## Classification of Queueing Systems

To define a queueing system precisely we need to specify (1) the arrival pattern, meaning the average rate of arrival of customers and the statistical distribution, (2) the service mechanism, that is, when service is available, how many servers there are, and how long service takes (usually specified by a distribution of service time), and (3) the queue discipline, that is, how a customer is selected for service from among the waiting customers, as for example, "first come, first served." Many interesting examples that illustrate some of the possibilities for the above characteristics appear in Cox and Smith [1], Bharucha-Reid [1], Feller [3], and Kendall [2], [3]. We must also specify whether there is interaction between the parts of the system, such as the arrival rate decreasing as potential customers are discouraged by a long queue.

One type of arrival pattern consists of single customers arriving at equally spaced instants, $t_1$ units of time apart. The average rate of arrival, $\alpha$, is then $1/t_1$ per unit of time. But, as we shall soon see, more suitable to a Markov chain application is an arrival pattern consisting of completely random arrivals. For completely random arrivals we assume that the number of arrivals in time interval $(t,t+\Delta t)$ is independent of $t$ and of the number of arrivals in any time interval not overlapping $(t,t+\Delta t)$. Hence, it is of no advantage to have information on the arrivals prior to time $t$ to predict the number of arrivals in $(t,t+\Delta t)$. (The above assumption is shown by experience to be closely satisfied in a number of problems, according to Cox and Smith [1] and Feller [1].) Suppose that the average rate of arrival is given by the constant $\alpha$, $0 < \alpha < \infty$, and suppose that the probability of more than one arrival in time interval $h$ becomes negligibly small as $h \to 0$, that is, the arrivals are isolated. Divide the interval $(t,t+\Delta t)$ into $n$ subintervals of equal length $\Delta t/n$, for $n = 1, 2, \ldots$, and let $A_{nk} = 1$ if an arrival occurs in the $k$th subinterval, $k = 1, 2, \ldots, n$, and equal zero otherwise. Then

$$S_n = \sum_{k=1}^{n} A_{nk}$$

is the number of arrivals in the interval $(t,t+\Delta t)$ for $n$ sufficiently large, under the assumption of isolated arrivals. Note that the expected number of arrivals in the interval is $\alpha\Delta t$. Thus,

$$P\{S_n = r\} = \frac{n!}{r!(n-r)!} p_n^r (1 - p_n)^{n-r}$$

since we have the Bernoulli case, where

$$p_n = P\{A_{nk} = 1\}$$

for $k = 1, 2, \ldots, n$, in accordance with the earlier assumption for intervals of equal length. Thus the expected value of $S_n$ is $n\ p_n$, and

$$n\ p_n = \alpha\Delta t \ ,$$

or,

$$p_n = \frac{\alpha\Delta t}{n} \ .$$

Then

$$P\{S_n = r\} = \frac{n!}{r!(n-r)!}\left(\frac{\alpha\Delta t}{n}\right)^r \left(1 - \frac{\alpha\Delta t}{n}\right)^{n-r}$$

$$= \frac{(\alpha\Delta t)^r}{r!}\left(1 - \frac{\alpha\Delta t}{n}\right)^n \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \ldots$$

$$\left(1 - \frac{r-1}{n}\right)\left(1 - \frac{\alpha\Delta t}{n}\right)^{-r} \ ,$$

and

$$\lim_{n\to\infty} P\{S_n = r\} = \frac{(\alpha\Delta t)^r}{r!} e^{-\alpha\Delta t} \qquad (39)$$

is the probability of r arrivals in the interval $(t, t+\Delta t)$. (The above result is essentially the Poisson Theorem of Loève [1].) In particular, the probability of no arrivals in the interval is given by $e^{-\alpha\Delta t}$, and hence the probability of one or more arrivals is $1 - e^{-\alpha\Delta t}$.

If customers arrive at times $t_1, t_2, \ldots$, where $0 < t_1 < t_2 < \ldots$, we let $u_n = t_{n+1} - t_n$ denote the time between the arrival of the nth and $(n + 1)$st customers. In the random case above, the arrival times $t_n$ are determined in a Poisson process (that is, randomly, where random is defined by the above assumptions on independence and isolation of arrivals) with parameter $\alpha$, the arrival rate. The inter-arrival times $u_n$, $n \geq 1$, have the negative-exponential distribution, with distribution function

$$A(u) = 1 - e^{-\alpha u}, \quad u \geq 0 ;$$

$$A(u) = 0 \qquad\qquad , \quad u < 0 ,$$

since $P\{u_n \leq u\}$ is the probability of an arrival in the time interval $(t_n, t_n + u)$, for $u \geq 0$, which was found to be $1 - e^{-\alpha u}$, and is zero for $u < 0$. In the random case, the random variables $u_n$, $n \geq 1$, will be independent.

A third arrival pattern, which actually includes both the regular and random patterns above, is called general independent. Here we suppose only that the random variables $u_n$, $n \geq 1$, are independent and have the same arbitrary distribution, given by the distribution function $A(u)$, where $A(u) = 0$ for $u < 0$.

Moving now to the service mechanism, we shall assume that the service times $v_n$, $n \geq 1$, are independent of each other and of the arrival pattern (and hence of the length of the queue), that the servers are identical if there is more than one, and that for all customers the service time has the same arbitrary distribution, given by the distribution function $B(v)$, where $B(v) = 0$ for $v < 0$. Also, let the average rate of customers leaving the queue after being served (no other way to leave the queue will be considered) be given by the constant $s\beta$, $0 < \beta < \infty$, where s is the constant number of servers, so the expected value of service time, v, is $1/\beta$. Note that the service time v refers only to the amount of time the customer is in contact with the server. This does not include the queueing time t that he waits in line before reaching the server. The customer's waiting time is $t + v$.

One type of service mechanism consists of regular departures, each customer being served for exactly the same length of time, $1/\beta$. Another consists of random service times, and exactly as before for the random arrival pattern, the service times $v_n$, $n \geq 1$, have the negative-exponential distribution, with distribution function

$$B(v) = 1 - e^{-\beta v}, \quad v \geq 0 ;$$

$$B(v) = 0 \qquad\quad , \quad v < 0 .$$

For general service times, $B(v)$ is not specified.

We have mentioned that the number of servers is given by the finite constant s, but we must also specify the availability of the servers. For example, there may be a probability of availability associated with one of the servers who performs other tasks that keep him away from the counter part of the time. In this paper we assume complete availability of the s servers.

The third and final element in specifying our queueing system is the queue discipline. This, of course, will have no effect on the number of people waiting for service, that is, on the queue length, but is important in studying the waiting time and the free periods of the individual servers. We shall deal only with "first come, first served," although, as is pointed out by Feller [3], it is by no means uncommon to have "last come, first served" or random choice as the rule. When there are $s > 1$ servers there are three principal ways in which customers may be assigned to servers on a "first come, first served" basis: (1) the customers are assigned to servers in strict rotation, regardless of the length of the queue for the individual server, (2) the customers select an arbitrary queue, usually the shortest, or (3) a single queue is formed and the customer at the head of the queue is served as soon as a server becomes free. (See Cox and Smith [1] for a more detailed discussion along these lines.) If we allow changes from queue to queue in the second case it does not differ from the third in total queue length, for then in both cases no server is idle until all customers are being served, but, of course, waiting times may be affected. In this paper changes will be allowed, for otherwise we would have s separate queues with variable arrival rates dependent on s queue sizes.

## The Imbedded Markov Chain

If the state of a stochastic system at time t is described by a random function $X(t)$, analogously to the Markov assumption for discrete random variables in Chapter II, we say that the stochastic process is Markovian if knowledge of the present value of $X(t)$ makes all information about the past history of the process irrelevant to a prediction of the future behavior of the process.

If the state of the process is measured by the queue size $q(t)$ (which includes those being served), and the arrivals are Poissonian and service times are negative exponential, we see that the above requirement for the process to be Markovian is satisfied. But, in general, the process measured by queue size alone is not Markovian. For example, if service times are regular and arrivals are Poissonian $q(t)$ and $v_i$, $i = 1, 2, \ldots, s$, the time already expended at time t on the customer being served by server i, are necessary to make the process Markovian. By including enough information in $X(t)$, it would seem that we could always make the process Markovian. As pointed out by Feller [1] this may be true, but the analysis of the queueing system will be less complicated if $q(t)$ will suffice.

For our queueing systems $X(t)$ will be an integer-valued step function, defined to be continuous from the right at the points of discontinuity. Let $\Omega_t$ denote the set whose elements are the functions with domain $(-\infty,t]$ and range the same as $X(\cdot)$, where $X(\cdot)$ represents the history of the system, with domain $(-\infty,\infty)$. For each t in $(-\infty,\infty)$, let $\Theta_t$ be a specified subset of $\Omega_t$, and let $\Pi$ be the set of those t for which $\Theta_t$ contains the contraction of $X(\cdot)$ to $(-\infty,t]$. For each

$t \varepsilon \Pi$ let $f_t$ be a specified functional with domain $\Theta_t$ and set

$$Y(t) = f_t\{X(\tau): \tau \leq t\} \quad .$$

If $\{\Theta_t, f_t: -\infty < t < \infty\}$ are chosen such that (1) $\Pi$ has no finite accumulation point, so its members can be written as $\ldots < t_{n-1} < t_n < t_{n+1} < \ldots$ , and (2) distribution $\{y_{n+1}|y_n, y_{n-1}, \ldots\}$ = distribution $\{y_{n+1}|Y_n\}$ for all n, where $y_m = Y(t_m)$, then the variables $\ldots, y_{n-1}, y_n, y_{n+1}, \ldots$, constitute an *imbedded Markov chain.*

For example, choosing $\Theta_t = \Omega_t$ for each integer t and $\Theta_t$ empty for all other t, $\Pi$ is the set of all integers. Then choosing $f_t \equiv 1$ we have an imbedded Markov chain. However, for our investigation of queueing systems we will want $f_t$ to be chosen so that Y(t) will be useful in describing the state of the queue.

The remainder of this chapter consists of applying the method of the imbedded Markov chain to two special queueing systems discussed by Foster [1] and Kendall [2], [3].

## The Random Arrival Queueing System

Consider now the queueing system with random arrivals, one server, and general service times satisfying the standard assumptions stated earlier. Let q denote the length of the queue and let the arrival and departure rates be $\alpha$ and $\beta$, respectively.

If the state of the system is described by $X(t) \equiv q(t)$ then the process is not necessarily Markovian (unless the service times are negative exponential), since knowledge of the expended service time is needed to make past history irrelevant. Define $X(\cdot)$ to be continuous from the right at its points of discontinuity and for membership in $\Theta_t$ require that $X(t) = X(t-) - 1$ (that is, q has just decreased by one), so $\Pi$ consists of the epochs of departure. For $t \in \Pi$ define

$$f_t\{X(\tau): \tau \leq t\} = X(t) ,$$

so $Y(t) = q$ is the number of customers left behind by a departing customer. Note that (1) and (2) of the definition are satisfied so we have an imbedded Markov chain. (The above technique fails if there are $s > 1$ servers since even at an epoch of departure there are $s - 1$ expended service times left unspecified.)

Define, for arbitrary n,

$$p_{ij} = P\{y_{n+1} = j | y_n = i\} ,$$

$i, j = 0, 1, 2, \ldots$, so $[p_{ij}]$ is the transition probability matrix for the imbedded Markov chain. Thus, by considering the queueing system only at the epochs of departure we have a Markov chain with a countable number of states and, since the service time has the same distribution for all customers, constant transition probabilities. Clearly

$$p_{ij} > 0 , \quad \sum_j p_{ij} = 1$$

for i, j = 0, 1, 2, ..., so we have a stochastic matrix. The form of $[p_{ij}]$ is

$$p_{oj} = k_j , \qquad j = 0, 1, 2, \ldots,$$

$$p_{ij} = k_{j+1-i} , \qquad i = 1, 2, \ldots, \qquad j \geq i - 1 ,$$

and

$$p_{ij} = 0$$

for all other i and j, where $k_r$ is the probability that there will be exactly r arrivals during a single service time. That is, the entries of $[p_{ij}]$ are

| i \ j | 0 | 1 | 2 | 3 | . |
|---|---|---|---|---|---|
| 0 | $k_o$ | $k_1$ | $k_2$ | $k_3$ | . |
| 1 | $k_o$ | $k_1$ | $k_2$ | $k_3$ | . |
| 2 | 0 | $k_o$ | $k_1$ | $k_2$ | . |
| 3 | 0 | 0 | $k_o$ | $k_1$ | . |
| . | . | . | . | . | . |

If B(v) is the service time distribution function then using (39),

$$k_r = \int_o^\infty \frac{(\alpha v)^r}{r!} e^{-\alpha v} \, dB(v)$$

for $r = 0, 1, \ldots$, and note that $k_r > 0$. Thus, $p_{ii} > 0$, $i = 0, 1, \ldots$, so all states are aperiodic and, by Theorem 18, $p_{ij}^{(n)} \to f_{ij}/\mu_{jj}$ as $n \to \infty$ for all i and j. In addition, from any state we may reach state zero with positive probability in a finite number of steps and may reach any state from state zero in one step with positive probability. By Theorem 6, the Markov chain is irreducible.

We now define the relative traffic intensity $\rho$ by

$$\rho = \frac{\alpha}{\beta} = \sum_{n=1}^{\infty} n\, k_n ,$$

the expected number of arrivals during a single service time. We would expect that for $\rho < 1$ the system would eventually settle into equilibrium operation, whereas for $p > 1$ the queue would become arbitrarily long. This conjecture, and the case $\rho = 1$, will now be investigated.

Suppose $\rho < 1$. Define

$$y_j = \frac{j}{1 - \rho}$$

for $j = 0, 1, \ldots$ . Then

$$\sum_{j=0}^{\infty} p_{oj}\, y_j = \sum_{j=0}^{\infty} \frac{k_j\, j}{1 - \rho} = \frac{\rho}{1 - \rho} < \infty ,$$

and, for $i \neq 0$ ,

$$\sum_{j=0}^{\infty} p_{ij}\, y_j = \sum_{j=i-1}^{\infty} \frac{K_{j+1-i}\, j}{1 - \rho} = \frac{1}{1 - \rho} [(i - 1) \sum_{j=0}^{\infty} k_j + \sum_{j=0}^{\infty} j\, k_j]$$

$$= \frac{1}{1 - \rho} (i - 1 + \rho) = y_i - 1 ,$$

and thus Theorem 28 indicates that the Markov chain is positive.

Suppose $\rho > 1$. Let

$$A(s) = \sum_{n=0}^{\infty} k_n s^n - s .$$

According to the treatment of generating functions in the proofs of Theorems 15 and 21, $A(s)$ is continuous on $[-1,1]$,

$$A'(s) = \sum_{n=1}^{\infty} n k_n s^{n-1} - 1$$

on $(-1,1)$, and

$$\lim_{s \to 1^-} A'(s) = \sum_{n=1}^{\infty} n k_n - 1 = \rho - 1 > 0 .$$

In addition, $A(0) = k_o > 0$ and $A(1) = 0$, and hence there exists an $s_o$, $0 < s_o < 1$, such that $A(s_o) = 0$. Define $y_i = s_o^i - 1$, $i = 1, 2, \ldots$ . Then

$$\sum_{j=1}^{\infty} p_{1j} y_j = \sum_{j=1}^{\infty} k_j y_j = \sum_{j=1}^{\infty} k_j (s_o^j - 1)$$

$$= A(s_o) + s_o - K_o - (1 - k_o) = s_o - 1 = y_1 ,$$

and for $i = 2, 3, \ldots,$

$$\sum_{j=1}^{\infty} p_{ij}\, y_j = \sum_{j=0}^{\infty} k_j\, y_{i+j-1} = \sum_{j=0}^{\infty} k_j\, (s_o^{\,i+j-1} - 1)$$

$$= s_o^{\,i-1}\, [A(s_o) + s_o] - 1 = s_o^i - 1 = y_i \ .$$

In addition, $\{y_i\}$ is non-zero and bounded, since $0 < s_o^i < 1$, $i = 1, 2, \ldots$, and thus by Theorem 26 the Markov chain is transient.

Finally, suppose $\rho = 1$. Define $y_j = j$, $j = 0, 1, 2, \ldots$ . Then

$$\sum_{j=0}^{\infty} p_{ij}\, y_j = \sum_{j=i-1}^{\infty} k_{j+1-i}\, j = i - 1 + \rho = i = y_i$$

for $i \neq 0$, and $y_i \to \infty$ as $i \to \infty$ . By Theorem 30 the Markov chain is persistent. From the structure of the matrix $[p_{ij}]$, it is clear that the expected passage times satisfy $\mu_{i,i-1} = \mu_{10}$ for $i \neq 0$ and $\mu_{io} = \mu_{i,i-1} + \mu_{i-1,0}$ for $i \geq 2$, since to pass from state i to state zero we must pass through state i - 1. Thus,

$$\mu_{io} = \mu_{10} + \mu_{i-1,0} = \ldots = i\, \mu_{10} \ ,$$

and thus

$$\sum_{j=1}^{\infty} p_{1j}\, \mu_{jo} = \mu_{10} \sum_{j=1}^{\infty} j\, k_j = \mu_{10}\, \rho = \mu_{10} \quad .$$

By Theorem 29, if the Markov chain is positive the above sum must equal

$\mu_{10} - 1$, and from the proof of that theorem $\mu_{10}$ is finite. Thus, the Markov chain is null.

To summarize the above results, according as $\rho$ is less than, equal to, or greater than one, the Markov chain is, respectively, positive, persistent null, or transient.

Note that for $\rho \geq 1$ the mean recurrence time $\mu_{oo} = \infty$, and hence the busy periods are of infinite expected length. In addition, $p_{ij}^{(n)} \to \frac{f_{ij}}{\mu_{jj}} = 0$ as $n \to \infty$ for all i and j, and thus for any fixed N, arbitrarily large, the probability approaches zero that the nth departing customer will leave the system with fewer than N customers in the queue. But, since $f_{io} > 0$ by Theorem 7, there is a positive probability that the counter will eventually become free. Actually, for $\rho = 1$ the Markov chain is persistent so, using (37), $f_{io} = 1$. That is, in this case the probability is one that the counter will eventually be free.

From Theorem 21, no stationary distribution exists for $\rho \geq 1$, and the positive sequence $\{1/\mu_{jj}\}$ gives the unique stationary distribution for $\rho < 1$, independent of the initial distribution. Thus, for $\rho < 1$, the probability that the queue is of length j after the nth departure tends to $1/\mu_{jj}$ as $n \to \infty$, independently of the initial state of the system.

## The Random Departure Queueing System

In this section we shall investigate a slightly more involved queueing system along the same lines as in the previous section. Let the queueing system have s servers, random departures (that is, negative-exponential service times), and the general independent arrival pattern.

We shall treat the second and third queue disciplines, described earlier, in which no server is idle until all customers are being served. The remaining queue discipline may be obtained by regarding each of the s queues as a single server queueing system with arrival rate $\alpha/s$, where $\alpha$ is the arrival rate for the entire queueing system. Let $\beta$ be the departure rate from each individual server.

If the state of the system is described by $X(t) = q(t)$, where q is the total number of customers being served or waiting at time t, then to make past history irrelevant to a prediction of future behavior we need to know the time since the last arrival (unless arrivals are random). Note that now, in contrast to the situation in the preceding section, expended service time at each of the s servers is irrelevant, since service times have the negative exponential distribution. Thus, for membership in $\Theta_t$ we require that $X(t) = X(t^-) + 1$ (that is, q has just increased by one), so $\Pi$ consists of the epochs of arrival. For $t \in \Pi$, define

$$f_t\{X(\tau): \tau \leq t\} = X(t) - 1 ,$$

so that $Y(t) = q - 1$ is the number of customers, waiting or being served, ahead of the newly arrived customer. Then we have an imbedded Markov chain. (Note that the length of the queue in the ordinary sense, the number of waiting customers, is $Q = \max\{q - s, 0\}$.)

Define, for arbitrary n,

$$p_{ij} = P\{y_{n+1} = j | y_n = i\} ,$$

$i, j = 0, 1, 2, \ldots$ . Then $[p_{ij}]$ is the transition probability matrix for the imbedded Markov chain. Thus, by considering the queueing system only at the epochs of arrival, we have a Markov chain with a countable number of states, and constant transition probabilities, since the inter-arrival time has the same distribution for all customers. To describe the form of the matrix $[p_{ij}]$, we note that $p_{ij} = 0$ for $j > i + 1$, since if the nth arrival encounters i customers in the queue, the (n+1)st can encounter no more than $i + 1$ customers. But for $j \leq i + 1$, $p_{ij} > 0$, since for negative exponential service times there is a positive probability for any number of the waiting customers to be served during the inter-arrival time. Thus $[p_{ij}]$ has the form

$$\begin{array}{c|cccccc} i \diagdown j & 0 & 1 & 2 & 3 & \cdot \\ \hline 0 & p & p & 0 & 0 & \cdot \\ 1 & p & p & p & 0 & \cdot \\ 2 & p & p & p & p & \cdot \\ 3 & p & p & p & p & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \quad , \qquad (40)$$

where p indicates a positive entry. The columns $s, s + 1, \ldots$ (note that the first column is column zero), represent the cases $j = s$, $s + 1, \ldots$, that is, the case in which all servers are occupied and have been throughout the inter-arrival time. Thus for $i \geq j - 1$ ,

$$p_{ij} = P\{i + 1 - j$$

customers are served during an inter-arrival time$\}$ .

The departure rate is sβ since all servers are occupied, and using (39),

$$k_r = \int_o^\infty \frac{(s\beta u)^r}{r!} e^{-s\beta u} \, dA(u) \ , \qquad r = 0,\ 1,\ 2,\ \ldots$$

where $k_r$ is the probability of exactly r departures in a single inter-arrival time, and A(u) is the inter-arrival time distribution function. Thus $p_{ij} = k_{i+1-j}$ for $j \geq s$ and $i \geq j - 1$. The columns j = 0, 1, ... s - 1, are more difficult to describe since the servers are not all occupied throughout the inter-arrival time, causing the departure rate to vary from sβ. However, noting that

$$\sum_{r=0}^{\infty} k_r = 1 \ ,$$

The matrix $[p_{ij}]$ has the form

$$\begin{array}{c|ccc|ccc}
{}_i \diagdown {}^j & 0 & \ldots & s-1 & s & s+1 & \cdot \\
\hline
0 & & 1 & & 0 & 0 & \cdot \\
\vdots & & \vdots & & \vdots & \vdots & \vdots \\
s-1 & & \alpha_o & & k_o & 0 & \cdot \\
s & & \alpha_1 & & k_1 & k_o & \cdot \\
s+1 & & \alpha_2 & & k_2 & k_1 & \cdot \\
\cdot & & \cdot & & \cdot & \cdot & \cdot \quad \cdot
\end{array} \qquad (41)$$

where the first s columns have been added together and

$$\alpha_i = \sum_{j=i+1}^{\infty} k_j \quad .$$

Using the form (40), we see that the imbedded Markov chain is aperiodic and irreducible.

We define the relative traffic intensity $\rho$ for this queueing system by

$$\rho = \frac{\alpha}{s\beta} = \frac{1}{\sum_{n=1}^{\infty} n\, k_n} , \tag{42}$$

the reciprocal of the expected number of departures during a single inter-arrival time when all s servers are occupied.

Suppose $\rho < 1$. As in the previous section for the function A(s), there exists a number $s_o$, $0 < s_o < 1$, such that

$$\sum_{n=0}^{\infty} k_n\, s_o^n = s_o , \tag{43}$$

since

$$\sum_{n=1}^{\infty} n\, k_n = \frac{1}{\rho} > 1 .$$

Let $y_i = s_o^{i-s+1}$, $i = s - 1, s, s + 1, \ldots$ . Then for $j \geq s$ ,

$$\sum_{i=0}^{\infty} y_i\, p_{ij} = \sum_{i=0}^{\infty} s_o^{j-s+i}\, k_i = s_o^{j-s} \sum_{i=0}^{\infty} k_i\, s_o^i = s_o^{j-s+1} = y_j ,$$

where $y_o, y_1, \ldots, y_{s-2}$ are arbitrary. For $j = s - 1, s - 2, \ldots, 1$, we solve successively the equations

$$y_j = \sum_{i=0}^{\infty} y_i \, p_{ij} \tag{44}$$

to determine $y_o, \ldots, y_{s-2}$, recalling the form (40). First

$$1 = y_{s-1} = \sum_{i=s-2}^{\infty} y_i \, p_{i,s-1} \; ,$$

and thus

$$y_{s-2} = \frac{1}{p_{s-2,s-1}} \left(1 - \sum_{i=s-1}^{\infty} s_o^{\,i-s+1} \, p_{i,s-1}\right) \; ,$$

and similarly for $y_{s-3}, \ldots, y_o$ . In addition,

$$\sum_{i=0}^{\infty} y_i \, p_{io} = \sum_{i=0}^{s-2} y_i \left(1 - \sum_{j=1}^{s-1} p_{ij}\right) + \sum_{i=0}^{\infty} s_o^i \left(1 - \sum_{j=1}^{s-1} p_{s-1+i,j} - \sum_{r=0}^{i} k_r\right)$$

$$= \sum_{i=0}^{s-2} y_i - \sum_{j=1}^{s-1} \sum_{i=0}^{s-2} y_i \, p_{ij} + \sum_{i=0}^{\infty} s_o^i \left(1 - \sum_{r=0}^{i} k_r\right)$$

$$- \sum_{j=1}^{s-1} \sum_{i=s-1}^{\infty} y_i \, p_{ij}$$

$$= \sum_{i=0}^{s-2} y_i - \sum_{j=1}^{s-1} \sum_{i=0}^{\infty} y_i \, p_{ij} + \frac{1}{1 - s_o} - \sum_{r=0}^{\infty} k_r \frac{s_o^r}{1 - s_o}$$

$$= y_o - y_{s-1} + \frac{1}{1 - s_o} - \frac{s_o}{1 - s_o} = y_o$$

using (43), (44), and $y_{s-1} = s_o^o = 1$. Thus, since

$$\sum_{i=0}^{\infty} |y_i| = \sum_{i=0}^{s-2} |y_i| + \frac{1}{1 - s_o} < \infty ,$$

the imbedded Markov chain is positive, by Theorem 27. Recalling (30) from the proof of Theorem 27, and using Theorem 21, it follows that

$$\{\frac{y_i}{\sum_{i=0}^{\infty} y_i}\} , \qquad i = 0, 1, \ldots ,$$

is the unique stationary distribution. Letting

$$\sum_{i=0}^{\infty} y_i = c ,$$

we have that, regardless of the initial distribution, the probability that a new arrival will find no *waiting* customers (that is, that $q \leq s$) tends to

$$\frac{1}{c}(y_o + y_1 + \ldots + y_{s-2} + 1 + s_o) = 1 - \frac{1}{c}(s_o^2 + s_o^3 + \ldots)$$

$$= 1 - \frac{1}{c}(\frac{s_o^2}{1 - s_o}) .$$

The probability of finding n waiting customers is given by

$$P\{Q = n\} = \frac{s_o^{n+1}}{c} , \quad \text{for } n \geq 1 ,$$

for the stationary distribution. Note that $\mu_{oo}$ is finite, and thus busy periods are of finite expected duration.

Now suppose $\rho \geq 1$. Then

$$\sum_{n=1}^{\infty} n\, k_n \leq 1 .$$

Let $y_i = 1$, $i = s - 1, s, s + 1, \ldots$ . Then

$$\sum_{i=0}^{\infty} y_i\, p_{ij} = \sum_{i=0}^{\infty} k_i = 1 = y_j$$

for $j \geq s$, where $y_o, \ldots, y_{s-2}$ are arbitrary. We now solve successively the equations

$$y_j = \sum_{i=0}^{\infty} y_i\, p_{ij} , \qquad 1 \leq j \leq s - 1 ,$$

to determine $y_{s-2}, \ldots, y_o$. This is possible as before, since for $1 \leq j \leq s - 1$,

$$\sum_{i=s-1}^{\infty} p_{ij} \leq \sum_{j=0}^{s-1} \sum_{i=s-1}^{\infty} p_{ij} = \sum_{i=0}^{\infty} \alpha_i = \sum_{n=1}^{\infty} n\, k_n \leq 1 .$$

In addition,

$$\sum_{i=0}^{\infty} y_i\, p_{io} = \sum_{i=0}^{s-2} y_i(1 - \sum_{j=1}^{s-1} p_{ij}) + \sum_{i=0}^{\infty} (\alpha_i - \sum_{j=1}^{s-1} p_{s-1+i,j})$$

$$= \sum_{i=0}^{s-2} y_i - \sum_{j=1}^{s-1} \sum_{i=0}^{\infty} y_i\, p_{ij} + \sum_{i=0}^{\infty} \alpha_i$$

$$= y_o - y_{s-1} + \sum_{n=1}^{\infty} n\, k_n \le y_o \ .$$

But $\sum_{i=0}^{\infty} |y_i|$ is infinite, and hence by Theorem 27 the imbedded Markov chain is null. Consider the possible solutions of the set of equations

$$y_i = \sum_{j=s}^{\infty} p_{ij}\, y_j \ , \tag{45}$$

$i = s, s + 1, \dots$ . Define

$$Y(z) = \sum_{n=s}^{\infty} y_n\, z^n$$

and

$$A(z) = \sum_{n=0}^{\infty} k_n\, z^n$$

for each z for which the series converge. Then, for $|z| < 1$ ,

$$\frac{1 - A(z)}{1 - z} = \frac{\sum_{n=0}^{\infty} k_n - \sum_{n=0}^{\infty} k_n\, z^n}{1 - z} = \sum_{n=1}^{\infty} k_n \sum_{j=0}^{n-1} z^j$$

$$= \sum_{n=0}^{\infty} z^n \sum_{i=n+1}^{\infty} k_i = \sum_{n=0}^{\infty} \alpha_n z^n \quad ,$$

and hence

$$\left|\frac{1 - A(z)}{1 - z}\right| = \left|\sum_{n=0}^{\infty} \alpha_n z_n\right| < \sum_{n=0}^{\infty} \alpha_n = \frac{1}{\rho} \leq 1 \quad .$$

Thus, since

$$A(z) - z = 1 - z \left[1 - \frac{1 - A(z)}{1 - z}\right] ,$$

$A(z) \nparallel z$ for $|z| < 1$ and

$$\frac{1 - z}{A(z) - z} = \frac{1}{1 - \sum_{n=0}^{\infty} \alpha_n z^n} = \sum_{n=0}^{\infty} a_n z^n , \qquad (46)$$

where the $a_n$ are positive numbers. Note that since

$$\sum_{n=0}^{\infty} \alpha_n = \frac{1}{\rho} \quad , \qquad (47)$$

The series $\sum_{n=0}^{\infty} a_n$ converges or diverges according as $\rho > 1$ or $\rho = 1$. In addition,

$$Y(z)[A(z)-z] = \sum_{n=s}^{\infty} \sum_{m=0}^{n-s} k_{n-m-s} y_{m+s} z^n - \sum_{n=s}^{\infty} \sum_{j=s}^{\infty} p_{nj} y_j z^{n+1}$$

$$= \sum_{n=s}^{\infty} \sum_{m=0}^{n-s} k_{n-m-s}\, y_{m+s}\, z^n - \sum_{n=s}^{\infty} \sum_{j=s}^{n+1} k_{n+1-j}\, y_j\, z^{n+1}$$

$$= k_o\, y_s\, z^s \quad .$$

Consequently, for all $|z| < 1$ for which the series for $Y(z)$ converges,

$$Y(z) = \frac{k_o\, y_s\, z^s}{A(z) - z} = k_o\, y_s\, z^s \, \frac{\sum_{n=0}^{\infty} a_n\, z^n}{1 - z} \tag{48}$$

$$= k_o\, y_s\, z^s \left(\sum_{n=0}^{\infty} z^n\right)\left(\sum_{n=0}^{\infty} a_n\, z^n\right)$$

$$= k_o\, y_s \sum_{n=0}^{\infty} z^{n+s} \sum_{i=0}^{n} a_i \; .$$

Note that we can construct a solution of the system (45) for arbitrary $y_s$, due to the form (41) of the matrix $[p_{ij}]$ for this queueing system, and $y_s \neq 0$ for every non-zero solution. Note also from form (41) that either $0 < y_s < y_{s+1} < \ldots$, or $0 > y_s > y_{s+1} > \ldots$, and

$$y_{s+1} = \frac{1 - k_1}{k_o}\, y_s \quad ,$$

$$y_{n+1} = \frac{1 - k_1}{k_o}\, y_n - \frac{1}{k_o} \sum_{j=s}^{n-1} k_{n-j+1}\, y_j \quad ,$$

for $n = s + 1, s + 2, \ldots$ . Thus, for $n = s, s + 1, \ldots,$

$$|y_{n+1}| < \frac{|y_n|}{k_o} < \ldots < \frac{|y_s|}{k_o^{n-s+1}} \quad ,$$

and thus

$$\sum_{n=s}^{\infty} |y_n z^n| \le \sum_{n=s}^{\infty} \frac{|y_s|}{k_o^{n-s}} |z^n| = |y_s z^s| \sum_{n=0}^{\infty} (\frac{z}{k_o})^n < \infty$$

for $|z| < k_o$ . Thus the series for $Y(z)$ converges at least in the interval $(-k_o, k_o)$ and there exist non-zero numbers $z$ for which the result (48) is valid. It follows that

$$y_{s+n} = k_o \, y_s \sum_{i=0}^{n} a_i \ , \qquad n = 1, 2, \ldots,$$

and hence there is a non-zero bounded solution of the system (45) if $\rho > 1$. No such solution exists if $\rho = 1$, since the series $\sum_{n=0}^{\infty} a_n$ converges or diverges according as $\rho > 0$ or $\rho = 1$, using (46) and (47). Hence, using Theorem 26, the imbedded Markov chain is persistent for $\rho = 1$, and transient for $\rho > 1$.

In summary, the imbedded Markov chain is positive for $\rho < 1$, persistent null for $\rho = 1$, and transient for $\rho > 1$. Notice from (42) that by choosing s sufficiently large $\rho$ can always be made less than one. The remarks in the concluding paragraphs of the preceding section hold also for the queueing system of the present section.

It should be emphasized, as by Feller [3], that the limiting probabilities in this and the previous section refer to the average of a large number of identical queueing systems and not to the fluctuations of an individual system.

# APPENDIX I

## SKETCH OF MARKOV CHAIN EXISTENCE RESULTS

Suppose that the set I, the space $\Omega$, the Borel field $\zeta$, and the sequence of random variables $\{x_n\}$ are as specified in Chapter II. For every $E_i$, $E_j \in I$ let numbers $p_i$ and $p_{ij}$ be given such that

$$p_i \geq 0 \ , \quad \sum_i p_i = 1$$

$$p_{ij} \geq 0 \ , \quad \sum_j p_{ij} = 1 \text{ for each } i \ .$$

We seek the existence of a probability measure P on $\zeta$ such that the sequence of random variables $\{x_n\}$ satisfies the Markov assumption, and the Markov chain $\{x_n\}$ has the prescribed initial distribution $\{p_i\}$ and transition matrix $[p_{ij}]$.

Following Chung [1] and Doob [1], define a set function $P_1$ on the set of cylinder sets as follows:

$$P_1\{s(E_{j_o}, \ldots, E_{j_n})\} = p_{j_o} \, p_{j_o j_1} \cdots p_{j_{n-1} j_n} \ .$$

Then by the Kolmogorov extension theorem (Kolmogorov [1], pp. 27-33; Loève [1], pp. 92-95), $P_1$ can be extended to be a probability measure P on $\zeta$.

We now show that the Markov assumption holds using this P and $\{x_n\}$ .

$$P\{x_n = E_k | x_{n_1} = E_{j_{n_1}}, \ldots, x_{n_r} = E_{j_{n_r}}\}$$

$$= \frac{\sum\limits_{v_{n_i}=j_{n_i}(i = 1,\ldots,r)} p_{v_o} p_{v_o v_1} \cdots p_{v_{n-1}k}}{\sum\limits_{v_{n_i}=j_{n_i}(i = 1,\ldots,r)} p_{v_o} p_{v_o v_1} \cdots p_{v_{n_r-1}v_{n_r}}}$$

$$= \frac{\sum\limits_{v_{n_i}=j_{n_i}(i = 1,\ldots,r)} p_{v_o} p_{v_o v_1} \cdots p_{v_{n_r-1}v_{n_r}}}{\sum\limits_{v_{n_i}=j_{n_i}(i = 1,\ldots,r)} p_{v_o} p_{v_o v_1} \cdots p_{v_{n_r-1}v_{n_r}}} \cdot$$

$$\frac{\left(\sum\limits_{v_{n_r}=j_{n_r}} p_{v_{n_r}v_{n_r+1}} \cdots p_{v_{n-1}k}\right)}{\sum\limits_{v_{n_i}=j_{n_i}(i = 1,\ldots,r)} p_{v_o} p_{v_o v_1} \cdots p_{v_{n_r-1}v_{n_r}}}$$

$$= \sum_{v_{n_r}=j_{n_r}} p_{v_{n_r}v_{n_r+1}} \cdots p_{v_{n-1}k}$$

$$= \frac{\sum\limits_{v_{n_r}=j_{n_r}} p_{v_o} p_{v_o v_1} \cdots p_{v_{n-1}k}}{\sum\limits_{v_{n_r}=j_{n_r}} p_{v_o} p_{v_o v_1} \cdots p_{v_{n_r-1}v_{n_r}}} = P\{x_n = E_k | x_{n_r} = E_{j_{n_r}}\}$$

where the notation is that of Chapter II, and the summations range over all subscripts that are not fixed by the equality statements.

Further,

$$P\{x_o = E_k\} = p_k$$

by definition, and

$$P\{x_{k+1} = E_j | x_k = E_i\} = \frac{\sum\limits_{v_o,\ldots,v_{k-1}} p_{v_o} p_{v_o v_1} \cdots p_{v_{k-1} i} \, p_{ij}}{\sum\limits_{v_o,\ldots v_{k-1}} p_{v_o} p_{v_o v_1} \cdots p_{v_{k-1} i}} = p_{ij} \quad ,$$

and thus the Markov chain $\{x_n\}$ has the prescribed initial distribution $\{p_i\}$ and transition matrix $[p_{ij}]$ .

# APPENDIX II

## PROOF OF NUMBER THEORY LEMMA

The following lemma from elementary number theory is stated on page 24.

Lemma. If $a_1$, $a_2$, ..., $a_n$ are n district positive integers with greatest common divisor $(a_1,\ldots,a_n) = 1$, then any integer $M > a_1 a_2 \ldots a_n$ can be represented in the form

$$M = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n ,$$

where the $x_i$ are positive integers.

Proof. The result is trivial for n = 1. We will suppose throughout that we have ordered the $a_i$'s so that $0 < a_1 < a_2 < \ldots < a_n$. Consider n = 2. Then

$$\frac{M - a_1}{a_2} > \frac{M - 2a_1}{a_2} > \ldots > \frac{M - a_2 a_1}{a_2} \qquad (49)$$

are $a_2$ distinct positive numbers. Suppose for positive integers j and k, $k < j \leq a_2$, and positive integer p, we have

$$\frac{M - ka_1}{a_2} - \frac{M - ja_1}{a_2} = p \quad .$$

Then $(j - k)a_1 = pa_2$, where $0 < j - k < a_2$, which contradicts $(a_1, a_2) = 1$. Hence the numbers (49) are of the form

$$C_i + \frac{k_i}{a_2} \quad ,$$

$C_i$ an integer, $1 \le k_i \le a_2$, where there are $a_2$ distinct $k_i$'s. Thus one $k_i$ must equal $a_2$. Then for positive integers k and p,

$$\frac{M - ka_1}{a_2} = p \quad ,$$

and

$$M = ka_1 + pa_2 \ .$$

Consider now $n \ge 3$. Note that

$$\sum_{k=2}^{n} a_k < (n-1)\, a_n \le a_{n-1}\, a_n \le a_2 \cdots a_n \quad ,$$

and thus $1 + a_2 + \ldots + a_n \le a_2 \ldots a_n$. It follows that

$$a_1(1 + a_2 + \ldots + a_n) \le a_1\, a_2 \ldots a_n \quad .$$

Now let $M > a_1\, a_2 \ldots a_n$ be chosen. Then $M \equiv r \pmod{a_1}$, $1 \le r \le a_1$. Consider $(a_2, \ldots, a_n) = d \ge 1$. Then by the Euclidean algorithm

$$d = y_2 a_2 + \dots + y_n a_n ,$$

where the $y_i$ are integers. But $(d, a_1) = 1$, and thus

$$1 = z_1 d + z_2 a_1 ,$$

where $z_1$ and $z_2$ are integers. Thus

$$M = z_1 Md + z_2 Ma_1$$

$$= z_1 M(y_2 a_2 + \dots + y_n a_n) + z_2 Ma_1 ,$$

and accordingly

$$z_1 M(y_2 a_2 + \dots + y_n a_n) \equiv r \pmod{a_1}$$

since $z_2 Ma_1 \equiv 0 \pmod{a_1}$ and $M \equiv r \pmod{a_1}$. Now

$$z_1 My_i \equiv r_i \pmod{a_1} ,$$

$i = 2,\dots,n$, where $1 \leq r_i \leq a_1$. Thus

$$r_2 a_2 + \dots + r_n a_n \equiv r \pmod{a_1} ,$$

and the $r_i$ are positive integers. But

$$r_2 a_2 + \ldots + r_n a_n \le a_1(a_2 + \ldots + a_n)$$

$$< a_1(1 + a_2 + \ldots + a_n)$$

$$\le a_1 a_2 \ldots a_n < M$$

so that

$$M - (r_2 a_2 + \ldots + r_n a_n) > 0$$

and is congruent to zero mod $a_1$. Thus,

$$M - (r_2 a_2 + \ldots + r_n a_n) = p a_1 \quad ,$$

with p a positive integer, and

$$M = p a_1 + r_2 a_2 + \ldots + r_n a_n \ .$$

# LITERATURE CITED

[1] T. M. Apostol, *Mathematical Analysis, A Modern Approach to Advanced Calculus.* Reading: Addison-Wesley, 1957.

[1] A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and Their Applications.* New York: McGraw-Hill, 1960.

[1] K. L. Chung, *Markov Chains with Stationary Transition Probabilities.* Berlin: Springer-Verlag, 1960.

[1] D. R. Cox and W. L. Smith, *Queues.* New York: Wiley, 1961.

[1] J. L. Doob, *Stochastic Processes.* New York: Wiley, 1953.

[1] W. Feller, *On the Theory of Stochastic Processes, with Particular Reference to Applications;* Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, edited by J. Neyman. Berkeley: University of California Press, 1949, pp. 403-432.

[2] W. Feller, *An Introduction to Probability Theory and Its Applications,* vol. 1, 2nd ed. New York: Wiley, 1957.

[3] W. Feller, *Chance Processes and Fluctuations;* Modern Mathematics for the Engineer, Second Series, edited by E. F. Beckenbach, Chapter 6. New York: McGraw-Hill, 1961.

[1] F. G. Foster, *On the Stochastic Matrices Associated with Certain Queueing Processes;* The Annals of Mathematical Statistics, vol. 24, 1953, pp. 355-360.

[1] Fréchet, *Recherches théoriques modernes sur le calcul des probabilités, II. Méthode des fonctions arbitraires. Théorie des événements en chaine dans le cas d'un nombre fini d'états possibles,* 2d ed. Paris: Gauthier-Villars, 1952.

[1] J. G. Kemeny and J. L. Snell, *Finite Markov Chains.* Princeton: Van Nostrand, 1960.

[1] D. G. Kendall, *On Non-dissipative Markoff Chains with an Enumerable Infinity of States;* Proceedings of the Cambridge Philosophical Society, vol. 47, 1951, pp. 633-634.

[2] D. G. Kendall, *Some Problems in the Theory of Queues;* Journal of the Royal Statistical Society, Series B, vol. 13, 1951, pp. 151-185.

[3] D. G. Kendall, *Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain;* The Annals of Mathematical Statistics, vol. 24, 1953, pp. 338-354.

[1] A. N. Kolmogorov, *Foundations of the Theory of Probability.* New York: Chelsea, 1956.

[1] N. Levinson, *An Elementary Proof of the Stationary Distribution for an Irreducible Markov Chain;* The American Mathematical Monthly, vol. 72, 1965, pp. 366-369.

[1] M. Loève, *Probability Theory,* 2nd ed. Princeton: Van Nostrand, 1960.

[1] M. Rosenblatt, *Random Processes.* New York: Oxford University Press, 1962.